University of Alberta

Edwin Hubble and the Interpretation of Red-Shifts in the Spectra of Extragalactic Nebulae

By

Mark A. Kahl

A thesis to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Arts.

In

History

Department of History and Classics

Edmonton, Alberta

Spring 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-82207-9

Canada

University of Alberta

Library Release Form

Name of Author: Mark A. Kahl

Title of Thesis: Edwin Hubble and the Interpretation of Red-Shifts in the Spectra of Extragalactic Nebulae

Degree: Master of Arts

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form without the author's prior written permission.

Mark A. Kahl

7244 Sindall Rd.

Baltimore, MD 21234

October 29, 2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Edwin Hubble and the Interpretations of Red-Shifts in the Spectra of Extragalactic Nebulae* submitted by Mark A. Kahl in partial fulfillment of the requirements for the degree of Master of Arts in History.

Dr. Robert Smith - Supervisor

Dr. David Johnson - Committee Chair

Dr. Andrew Ede - Committee Member

Dr. Douglas Hube - Committee Member

October 25, 2002

To my parents.

Abstract

The observed red-shifts in the spectra of extragalactic nebulae have been interpreted as evidence of the expansion of the universe. For his extensive research on extragalactic nebulae and the formulation of the velocity-distance relation in 1929, Edwin Hubble is often credited with this discovery. However, in the early- to mid-twentieth century, not all astronomers accepted this proposition. At times, even Hubble appeared to have his doubts and he publicly refrained from drawing any definite conclusions. I have therefore examined Hubble's researches with the intent of uncovering his preferred interpretation of red-shifts. I argue that, despite his expressed methodology in which he gave primacy to observations, the weight of the evidence indicates that here he followed the dictates of theoretical presuppositions. The thesis provides an important case study on the acquisition of scientific knowledge as well as on how a leading scientist viewed the relationship between theory and observation.

Acknowledgements

I would like to thank my advisor, Dr. Robert W. Smith, for his help and guidance throughout the production of this thesis. I would also like to thank the Huntington Library for providing me with copies of some primary documents, my family and friends for their support, especially Amy Piotrowski, as well as fellow graduate student, Rebecca Adell, for her advice and time generously spent proof reading. Finally, I would like to thank Dr. Joseph Tatarewicz for introducing me to the field of the History of Science.

Table of Contents

Introduction.....1

Chapter 1: Background Information.....10

a. Spectral Analysis and Explanation of Red-Shifts.....10

b. History of the Velocity-Distance Relation.....12

Chapter 2: Cosmological Models and Interpretations of Red-Shifts.....22

a. Hubble's Evaluation of the Significance of Red-Shifts.....26

Chapter 3: Bringing the Pieces Together.....43

Conclusion.....64

Bibliography.....88

Introduction

> The history of astronomy is a history of receding horizons…At first the explorations (of space) were confined to the realm of the planets, then they spread through the realm of the stars, and finally they penetrated into the realm of the nebulae…The distances were the essential data. Until they were found, no progress was possible.[1]

This is a statement made by the eminent twentieth-century astronomer, Edwin Hubble, whose contributions to the study of the nebulae[2] are unparalleled. Hubble was born on 20 November 1889 in Marshfield, Missouri, but he grew up in Kentucky and attended high school in Chicago. In 1906 he was awarded a scholarship to the University of Chicago where he studied mathematics and astronomy. Hubble was also a talented athlete and earned a college letter in track, boxing, and basketball. In 1910 he received his Bachelor of Science and was awarded a Rhodes scholarship. For the next two years Hubble studied Roman and English law at Queen's College, Oxford. He returned to Kentucky in 1913, and his love of astronomy drew him back to the University of Chicago in 1914 to pursue postgraduate work leading to his doctoral degree. In 1917 Hubble

---

1 Hubble, 1936b, p. 21-22.

2 Here, an explanation of terminology will be useful for the reader. The use of the word nebulae should not be interpreted by what we consider nebulae to mean today, clouds of interstellar gas and dust. Rather, it should be interpreted as Hubble intended to use the term. In *The Realm of the Nebulae*, written in 1936, Hubble explains, "Today, the term nebula is used for two quite different kinds of astronomical bodies. On the one hand are the clouds of dust and gas, numbering a few score in all, which are scattered among the stars of the galactic system. These have been called galactic nebulae. On the other hand are the remaining objects, numbering many millions, which are now recognized as independent stellar systems scattered through space beyond the limits of the galactic system. These have been called extragalactic nebulae. The nomenclature is followed in this book with the exception that, since extragalactic nebulae are mentioned so frequently, the adjective will be dropped. Therefore the term nebulae will refer to extragalactic nebulae alone, unless otherwise specified." Hubble, 1936b, p. 16-17. This thesis will use terminology similar to that which Hubble used. Thus, when the term nebulae is used, the reader can infer that what is meant is extragalactic nebulae, or what we would refer to today as external galaxies.

obtained his Ph. D. for his thesis "Photographic Investigations of Faint Nebulae." He enlisted in the army to serve in the First World War soon thereafter.

Hubble returned to the U.S. in the summer of 1919 and joined the staff of the Mount Wilson Observatory in Pasadena, California. The director of the observatory, George Ellery Hale, had offered a position to Hubble in 1917, but Hubble turned it down to serve in the military. By the early twentieth century, Mount Wilson was regarded as the world's premier astrophysical observatory and it attracted many great astronomers and physicists. In 1917 the observatory became the home of the Hooker 100-inch telescope, which reigned as the world's largest telescope until the Hale 200-inch telescope was completed in 1948 for the Palomar Observatory. Working with this and other powerful telescopes at Mount Wilson, Hubble was able to make observations that were not possible at other observatories at the time. Thus, Mount Wilson provided Hubble with the ideal location and instruments to conduct his research on extragalactic nebulae.[3]

Until the early 1920's, very little was known about the nature of the nebulae. For over a century, it had been debated whether these faint patches of light were distant star systems, or if they were comprised of some luminous fluid. Also, if the nebulae were systems of stars, were they part of our galaxy, or were they galaxies in their own right?[4]

As mentioned above, the essential data required to settle the issue was an accurate determination of nebular distances. In 1923, Hubble discovered what astronomers considered the most reliable measure of distance for nebulae when he first detected the

---

[3] For more accounts of Hubble's personal life, see Christianson, 1995; Mayall, 1970; and Jones, 1989.

presence of a Cepheid variable in the Andromeda Nebula. By calculating the variable star's period and utilizing the period-luminosity curve, Hubble was able to estimate the nebula's distance to be about 300,000 parsecs, or approximately one million light-years. This placed the Nebula at a distance much farther than even the greatest estimates of the size of the Milky Way, and therefore provided conclusive proof that the nebulae were in fact independent galaxies or "island universes." This discovery alone would have been sufficient to secure Hubble's name in history as one of the world's greatest astronomers; however, his name would later become associated with an even greater discovery, the expansion of the universe.

When astronomers began to study nebular spectra in the early twentieth-century, they noticed that the wavelengths of the light emitted by distant nebulae were shifted to the red. That is, their light waves were red-shifted or apparently elongated. In 1929, aided by his Mount Wilson collaborator Milton Humason, Hubble formulated what would later become known as a law of red-shifts, which states that a linear relationship exists between the spectral red-shift and the distance of a nebula. This not only provided astronomers with another method of determining distance, but, as we shall see, was also of great cosmological significance.

The discovery of red-shifts was hardly a new phenomenon. Red-shifts were apparent in the light from stars, albeit to a much smaller degree than for the nebulae, and astronomers were familiar with them. In this case, red-shifts were interpreted as Doppler- or velocity-shifts caused by radial motion, that is, motion in the line of sight. Many astronomers in the 1930's typically applied this same interpretation to the red-shifts of

---

[4] Much has been written on the debate over the nature of the nebulae. See for instance Smith, 1982;

nebulae, and believed this was evidence that the universe was expanding. However, although the red-shifts in the spectra of the nebulae apparently behaved as velocity-shifts, interpreting them in this way was not regarded as straightforward. Many astronomers, including Hubble, acknowledged that the data obtained from observations did not provide sufficient evidence for this interpretation. In fact, as Hubble pointed out, the data could be manipulated to allow for alternative interpretations just as easily. Depending on the interpretation adopted, one could obtain a vastly different picture of the universe. Therefore, the interpretation of red-shifts was of fundamental importance to astronomers deciding on a cosmological model that explained the nature of the universe.

Since the proper interpretation of red-shifts was so important to cosmology, Hubble proceeded with caution, especially in his earlier published papers where he was mostly concerned with establishing the velocity-distance relation, and avoided judgments about possible interpretations. In a paper co-authored with his collaborator Milton Humason from 1931, Hubble stated that, "the writers are constrained to describe the 'apparent velocity-displacements' without venturing on the interpretation and its cosmological significance."[5] Throughout his published papers and books on the topic, he was always careful to point out that his use of the term velocity should be taken as "apparent velocity." In describing his research, Hubble wrote that he chose to emphasize the empirical aspects of his work, rather than partake in theoretical interpretations. Only later, when his velocity-distance relation gained general acceptance did he begin to deal in print with theoretical implications. Even then, he did not explicitly side with a particular cosmological model. Instead, Hubble identified the inconsistencies between

---

Hoskin, 1976.

observations and theories, and pointed out the possibility of errors in the data, frequently reminding his readers that a definite conclusion would require the use of larger telescopes and more observations.

The ambiguity in terms of views expressed in Hubble's published works regarding the interpretation of red-shifts is also apparent in his personal correspondence with colleagues. For instance, in a letter from the astronomer Nicholas Mayall to Hubble dated 6 March 1934, Mayall states that, "Moore is convinced of the necessity for including some kind of a correction for red-shifts, but he holds the same opinion of the interpretation of red-shifts as a velocity that you have already expressed in your velocity-distance paper."[6] Hubble produced two main papers on the topic of the velocity-distance relation prior to this letter (the first was written in 1929 and the second was co-authored with Humason in 1931). In both, however, he does not discuss interpretations in detail. For example, in the 1931 paper interpretations are only mentioned to qualify the use of the term velocity. Since the interpretation of actual motion could not confidently be accepted at that point, the authors explained that the term velocity was being used in the sense of "apparent" velocity. In his 1929 paper Hubble devotes less than a paragraph to interpretation, stating:

> The outstanding feature, however, is the possibility that the velocity-distance relation may represent the de Sitter effect, and hence numerical data may be introduced into discussions of the general curvature of space. In the de Sitter cosmology, displacements of the spectra arise from two

---

5 Hubble and Humason, 1931, p.80.
6 Mayall to Hubble, 1934.

> sources, an apparent slowing down of atomic vibrations and a general tendency of material particles to scatter.[7]

In addition, in a later letter from 1937, commenting on Hubble's book, *The Realm of the Nebulae*, Mayall says, "It is perhaps unnecessary to mention how pleased are some of the people here to note the way that your interpretation of the nature of the red-shift casts doubt upon the validity of theories of the expanding universe."[8] In his book, as in all of Hubble's published works, he presents a conservative and cautious view towards the interpretation of red-shifts. As I shall argue, although he acknowledges the discrepancies introduced by the interpretation of actual motion in his book, he also regards it as the only permissible interpretation.

This lack of clarity with respect to the position Hubble took in his writings regarding the interpretation of red-shifts has led historians to draw contradictory conclusions about Hubble's cosmological views. Some historians, such as Stephen Brush, view this ambiguity as evidence that Hubble did not believe that red-shifts were caused by actual motion, and argue that Hubble doubted the expanding model of the universe. Brush contends that:

> Hubble consistently avoided any definite statement that the red-shifts of distant nebulae are due to motion [though he admitted that this interpretation did apply to nearby ones], or that the universe is actually expanding, and seemed to doubt the validity of this conclusion. Thus one may say that Hubble 'discovered the expanding universe' in the same sense that Max Planck 'discovered the quantum': he established an empirical formula that seemed to imply the theory and indeed led

---

7 Hubble, 1929, p. 173.

> others to adopt it [and later to assume that he must have adopted it himself]- yet he drew back from explicitly advocating it as a true statement about the world, and on some occasions even suggested that it was false.[9]

On the other hand, other historians disagree with this interpretation of Hubble's work. For example, Norriss Hetherington contends that, "Reading Hubble's papers from the 1930's, it is all too easy now to conclude that he doubted the expanding model because his evidence was against it."[10] Hetherington argues that, despite the observations, Hubble believed in an expanding universe because his philosophical values dictated allegiance to it.[11]

This thesis will add to the existing literature on Hubble's career in a variety of ways. To begin with, those historians who argue that Hubble did not believe red-shifts were velocity-shifts appear to be too willing to accept his writings at face value. Rather than indicating Hubble's reluctance to accept an expanding model, I will argue that Hubble's ambiguity reveals the uncertainty held by most astronomers in the 1930s and 1940s – uncertainty over the nature of red-shifts and, by extension, the nature of the universe. Hubble's ambiguity is important in and of itself; it points to the fact that if one contextualizes his writings historically, it becomes apparent that a disjuncture exists between his methodological stipulations and the philosophical preconceptions that dominated his work.

---

8 Mayall to Hubble, 1937.
9 Brush, 2001, p. 162. For examples of similar arguments see North, 1990b; and Ellis, 1990.
10 Hetherington, 1996, p. 173.
11 See Hetherington, 1982.

In this thesis, I will provide an examination or reexamination of Hubble's work on spectral red-shifts with the intent of uncovering his actual interpretation of the phenomenon, and his reasons for accepting this interpretation. Like Hetherington, I believe that philosophical values greatly influenced Hubble's evaluation and interpretation of his data. However, I aim to proceed further and actually identify those fundamental beliefs and preconceived values, and examine how they in fact influenced his decisions and arguments.

I will begin the first chapter with a brief overview of the history of the investigation of red-shifts. This will also provide an opportunity to introduce some of the researchers involved and the techniques they exploited for obtaining data. Since Hubble did not seem interested in interpreting his findings at the start of his investigations, except for his claim that he might have evidence for the de Sitter effect, I will place emphasis on the establishment of a relation between red-shifts and distance to nebulae. This discussion will conclude with Hubble's formulation of the velocity-distance relation.

In the second chapter of this thesis, I will deal with how the researchers analyzed the data, and the implications of their interpretations. After the effects of red-shifts were established in the form of a widely accepted law, Hubble began to discuss the various interpretations and implications suggested by the observations. Although he remained careful not to over-commit himself, his position can in fact be inferred from his discussions. Since his views, despite his claims, were so strongly influenced by theoretical presuppositions, it will be helpful to have a short discussion of theories and cosmological models that were current at the time. For instance, a basic discussion of general relativity, the various interpretations of red shifts, and the theories of men such as

Zwicky, Milne, Lemaître, de Sitter, and Eddington will be laid out before looking at what Hubble thought about the topics. This will lead into a discussion of how Hubble, with the help of others such as R. C. Tolman for example, described his evaluation of the data in terms of theory, and vice versa. An important point here is that although Hubble seemed to remain ambiguous about his interpretation, there are clues to his privately held views in the way in which he presented and evaluated the data.

In the last chapter, I will draw these various strands together to demonstrate Hubble's preferred interpretation of red-shifts. At the same time, I will argue that this preference was not founded on the basis of observation alone, but rather on the dictates of theory. The reasons he accepted this view underline the fact that he valued theory and his beliefs, even when they seemed contrary to observations, and reveal that in practice he followed a quite different set of principles from those expressed in his writings.

As one of history's greatest astronomers, and the leading authority in regards to red-shifts and extra-galactic nebulae in the period between the early 1920s and early 1950s, Hubble's beliefs on this immensely important cosmological matter are crucial indicators for understanding how top scientists of the early twentieth century viewed our universe. The rationale he offered for his beliefs, and the ways in which he evaluated data and presented arguments, are good indications of how scientists understood the relationship between theory and observations, as well as how scientific knowledge gained acceptance in a time when cosmology was becoming a respected scientific discipline in its own right.

Chapter 1: Background Information

Spectral Analysis and Explanation of Red-Shifts:

By the early 1920's, astronomers had long accepted that visible light consists of different wavelengths, which are crudely distinguished by the eye as colors. Long waves are seen as red, and the shorter waves are blue or violet. When a beam of light is passed through a prism, the individual rays are refracted to different degrees and the colors are spread out in an ordered sequence called a spectrum. The amount of refraction is dependent upon wavelength.

By this time, three types of spectra had been identified. The first is emitted by incandescent solids, such as electric light filaments. Their spectrum is continuous from red to violet, and beyond in each direction, and is therefore known as a continuous spectrum. The relative abundance of the colors in this spectrum measures the temperature of the light source.

The second type of spectrum is emitted by an incandescent gas, such as a neon light, which radiates only certain colors. This type of spectrum is known as an emission or bright-line spectrum, because it is comprised of a pattern of isolated lines separated by dark gaps. Similar to a fingerprint, the pattern of colors in this type of spectrum is a unique characteristic of the particular gas involved, and an unknown gas can be identified by its spectrum alone. Thus, an emission spectrum from a distant light source provides an observer with clues to the source's chemical composition.

The third type of spectrum is called an absorption spectrum, and consists of a continuous background that is interrupted by dark spaces or lines. A spectrum of this kind

results from an incandescent solid, or similar source that emits a continuous spectrum, being surrounded by a cooler gas. In this instance, the gas will absorb from the continuous spectrum those colors that it would emit itself if it were incandescent. An example of this occurrence is a star with its surrounding atmosphere. The dark lines in this type of spectrum represent the colors that are absorbed by the surrounding gas. Knowing which colors the surrounding gas absorbs leads to the identification of the gases in the stellar atmosphere. The distribution of intensity in the continuous part of an absorption spectrum determines the temperature of the star.

Since absorption spectra provide such crucial information about distant light sources, their analysis had become extremely important to astronomers by the 1920s. In *Red-Shifts in the Spectra of Nebulae*, written in 1934, Hubble stated:

> The study of absorption spectra is the dominating feature of modern astronomy. Absorption spectra furnish an astonishing amount of information concerning the physical condition of stars and even of planets and of nebulae. Either directly or indirectly they indicate surface temperatures of stars, surface luminosities, total luminosities, distances, [and] velocities in the line of sight.[12]

Radiation from the Sun produces an absorption spectrum. Absorption lines in the spectrum of the Sun indicate, for example, the presence of hydrogen and iron vapors in it's atmosphere. The most conspicuous absorption lines are the calcium lines near the violet end of its spectrum. By the 1920s these were known as the H and K lines. The spectra of what Hubble called extragalactic nebulae, what we now know as galaxies, are

---

[12] Hubble, 1934a, p. 6. Hubble's discussions on spectrum analysis are present in a number of his writings and remained consistent over the years. See for instance *The Observational Approach to Cosmology* (Hubble, 1937a) or *Red-Shifts in the Spectra of Nebulae* (Hubble, 1934a).

very similar to that of the Sun. Even though the light of the nebula is extremely faint, the H and K lines of calcium and lines corresponding to hydrogen and iron can often be easily identified, provided the nebula is not too faint for lines to be detected.

Although the spectra of the Sun and many extragalactic nebulae are similar, there is a significant characteristic of nebular spectra that distinguishes the two. In the vast majority of spectra of nebulae, absorption lines are displaced towards the red end of the spectrum. These displacements are known as red-shifts, and they are observed in all extragalactic nebulae except a few of those closest to our own Galaxy. Each line in the spectrum is shifted by a certain constant fraction of its normal wavelength λ. The amount of red-shift varies from one nebula to the next, and is represented by the term $d\lambda/\lambda$. On average, red-shifts increase with the apparent faintness of nebulae. Since apparent faintness is interpreted as a measure of distance, it can be stated that red-shifts increase with distance. In the opinion of his colleagues, by the early 1930's, Hubble established that this relation is approximately linear and can be represented as, $d\lambda/\lambda = k \times r$ (red-shift = constant x distance). This relation is called the law of red-shifts.

History of the Velocity-Distance Relation:

Around 1876, William Huggins and Henry Draper obtained the first photographic spectra of a nebula; however, little progress was made in the field until the turn of the twentieth century.[13] Some of the most significant early research on the analysis of nebular spectra was carried out by Vesto Melvin Slipher at the Lowell Observatory in Flagstaff, Arizona. The wealthy American, Percival Lowell, constructed this observatory in 1893-4.

Slipher joined the observatory's staff in 1901, and began working with a powerful new spectrograph in the spring of 1902. A few years later, Lowell suggested to Slipher that the latter should try to measure the rotation of the Andromeda Nebula. Many, including Slipher and Lowell, considered spiral nebulae, like the Andromeda Nebula, to be protostellar systems; and Slipher hoped to obtain clues about the origin of the solar system by studying the Nebula.[14]

In the process of determining the rotation of the Nebula, Slipher made an even greater discovery. On 17 October 1912, Slipher photographed the spectrum of the Andromeda Nebula, and for the first time measured its radial velocity. The result he obtained was –300 km/s, which indicated that the Andromeda Nebula was approaching our own Galaxy with a larger velocity than had been previously observed for any celestial object.[15] This was the start of a research campaign in which Slipher basically worked alone in the early years. By 1917, Slipher had collected the spectra of fifteen spiral nebulae[16], and by 1922, he had more than forty. In comparison, other observers had only assembled a half-dozen spectra from spiral nebulae.

Although astronomers had measured radial velocities via spectral shifts for stars, the measures of spectral shifts for nebulae contained some curious characteristics. The first was the enormous size of the measured radial velocities. For example, Slipher measured the velocities of NGC 4565 and 4594 to be +1000 and +1100 km/s

---

13 Fehrenbach, 1984, p.177.

14 Smith, 1990, p. 54.

15 Slipher, 1914(AAS), p. 99.

16 There is a list of these 15 nebulae and their velocities in Slipher, 1914(AAS), p. 100. See also Hoyt, 1980, and Osterbrock, 1990.

respectively.[17] As seen in these measures, another curious aspect of the data was that the vast majority of measurements were positive, that is, their light was shifted to the red. In 1923, the British astronomer Arthur Eddington commented on the measured red-shifts as follows:

> One of the most perplexing problems of cosmology is the great speed of the spiral nebulae. Their radial velocities average about 600 kilometers per second and there is a great preponderance of velocities of recession from the solar system. It is usually supposed that these are the most remote objects known (though this view is opposed by some authorities), so that here if anywhere we might look for effects due to general properties of the world… The great preponderance of positive (receding) velocities is very striking; but the lack of observations of southern nebulae is unfortunate, and forbids a final conclusion. Even if these also show a preponderance of receding velocities the cosmological difficulty is not entirely removed… It will be seen that two nebulae (including the great Andromeda Nebula) are approaching with rather high velocity and these velocities happen to be exceptionally well determined.[18]

Assuming that our own stellar system is a nebula, a proposition that was contested until 1924, attempts were made to derive the solar motion with respect to the other nebulae using the same method that determined the motion of the Sun with respect to the

---

17 Ibid.

18 Sciama points out that Eddington's words are a reminder that at the time it had not yet been definitely established that spiral galaxies were actually outside the Milky Way. Proof of this did not come until Hubble's discoveries the following year, 1924. In addition, the discovery of the Milky Way's rotation in 1926-1927 also shed further light on Slipher's velocities, and served to satisfy Eddington's concern about the rapidly approaching nebulae. "The Sun's velocity around the center of the Milky Way is about 250 km/s. Other objects in the Milky Way are also moving around its center, so that their radial velocity relative to the Sun is much less than 250 km/s. But objects outside the Milky Way do not share its rotation, so that their measured velocities must be corrected for the Sun's motion if we want to know their velocity relative to the Milky Way as a whole. When this correction was made the rapidly approaching galaxies which so worried Eddington were much slowed down…Thus the first velocity measured by Slipher, which

other stars. From 1916 onward attempts were made by Truman of the Iowa State Observatory and others, but the initial results were problematic in Hubble's opinion. This was because the measured positive velocities of the spiral nebulae, indicating general recession, proved intractable. The historian of science J. D. North maintains that "Truman and others obtained a value for this motion with respect to neighboring nebulae, but as information accumulated it became increasingly clear that the recorded nebular velocities could be accounted for only if, in addition to the solar velocity, there were some further systematic effects."[19] So, in 1918, as an attempt to "save the phenomena", Carl Wilhelm Wirtz introduced a constant K term into the equations that was supposed to represent a constant velocity to be subtracted from the apparent nebular velocity before evaluation of the solar motion.[20]

According to Hubble, this served to improve the calculations, but not enough to render them acceptable. However, Hubble believed this procedure led astronomers to suggest that the K term, which was supposed to be constant for all nebulae, should be

---

at the time seemed startling large, is actually a poor guide to the surprises that were to come." Sciama, 1971, p. 43.

[19] North, 1990a, p. 142.

[20] Hubble, 1953, p. 658. *As R.W. Smith points out, the method of analysis used in initial attempts to discern the solar motion with respect to spiral nebulae was straightforward and amounted to solving a set of equations that were well known by astronomers. In 1916, Truman of the Iowa State Observatory, and Young and Harper of the Dominion Observatory attempted such calculations; but as Hertzsprung, who made similar attempts prior to these men, stated uncertainties in their calculations made their results unreliable. Smith also mentions that G.F. Paddock introduced the K term in 1916, before Wirtz made use of it. However, Hubble does not mention Paddock in his account. A reason for this may have been that Paddock considered the term to be "merely a manifestation of an insufficient number of observations", and as more spirals were observed the term would disappear. Regardless, after Paddock's work, a K term was used in all subsequent attempts to calculate the solar motion from velocities of nebulae. See Smith, 1982, p. 166-168. By utilizing the recorded red-shifts and the observed diameters of nebulae, Wirtz attempted to establish a velocity-distance relation based on the size of red-shifts and a logarithm of diameters. However, without knowing the actual size of a nebula, estimating the distance based on its diameter is questionable. For more on Wirtz's use of the K term and attempts at establishing an early relation between red-shifts and distance see Seitter and Duerbeck, 1990; Wirtz, 1924; and North, 1990a.

replaced with a term that varied with the distances of individual nebulae.[21] A variable K term, a distance effect, was also suggested by de Sitter's studies of relativistic cosmology, which predicted red-shifts to increase with distance.[22] The early attempts to establish a variable K term were also unsuccessful. As Hubble recalled, the problem was not with the measured radial velocities, but rather in the means of estimating distances. In *Red-Shifts in the Spectra of Nebulae*, written in 1934, Hubble explained the difficulties he saw with the procedure as follows:

> The problem then consisted first in estimating this mean distance, and second in estimating relative distances of the individual nebulae in order to correlate them with the velocities. Unfortunately little was known concerning nebular distances, and the only obvious criteria, namely apparent luminosities and diameters, exhibited such large inherent dispersions as compared with the limited range of the few red-shifts available that the distance effect was more or less completely masked.[23]

At this point, Hubble reckoned, the solution seemed to lie in establishing an accurate determination of distance, and this is precisely what Hubble's work of 1929 offered. Essentially, Hubble's method for determining a criterion of distance followed a

---

21 Hubble, 1953, p. 658-659.

22 Hubble, 1934a, p. 7. De Sitter's predicted relation between red-shifts and distance was initially not meant to imply actual motion. In general, the "de Sitter effect," was used to describe the lowering of the frequency of light from distant sources as it traveled through space. North states that the effect was explained as a slowing down of time at great distances from the observer, manifesting itself as an apparent slowing down of atomic vibrations, and hence as a spectral shift. Therefore, the "de Sitter effect," by itself, could not be interpreted as indicating a general recession of the nebulae. However, as North points out, by the 1920's, such a recession became associated with de Sitter's theory. Eddington was one of the first to recognize the astronomical significance of de Sitter's theory, and he was aware of the two-fold explanation of spectral shifts it provided. Eddington maintained that, in the de Sitter world, particles would tend to scatter unless their mutual gravitation was sufficient to overcome this tendency. Since the density of de Sitter's universe was approximately zero, there would not be enough attraction and the particles would scatter under cosmical repulsion. See North, 1990a.

step-by-step procedure, which consisted of using a collection of distances for nearby nebulae to calibrate a successive method that would lead to greater distances. This procedure was based on the simple principles that if the intrinsic luminosity (or candle-power) of an object was known, then the apparent faintness of the object indicates distance. Likewise, if the distance of the object is known, then the apparent faintness indicates the intrinsic luminosity.[24]

The first collection of distances that Hubble deemed reliable was derived from investigations of certain stars within our own galactic system. In 1923, Hubble began to study novae in the Andromeda nebula, with the aim of accurately establishing its distance. As the historian of science, R.W. Smith states:

> Hubble expected to determine with higher precision the mean apparent brightness of novae in the nebula. Comparing the apparent brightness of novae in the Andromeda Nebula with the absolute brightness of novae in the Galaxy and employing the principle of the uniformity of nature- according to which novae in the nebula should be, on the average, as bright as those in the Galaxy- he hoped to estimate its distance with more certainty.[25]

In October 1923, Hubble discovered what at first he believed to be a nova, but later turned out to be a Cepheid variable.[26] Once Hubble realized what he had discovered, he calculated the period for the star's light variations. He then calculated its absolute

---

23 Hubble, 1934a, p.7-8.

24 Hubble, 1929, p. 168-169. See also Osterbrock, 1990.

25 Smith, 1990, p. 55.

26 Smith, 1990, p.56. Smith points out that Hubble's observing program was designed to discover novae, so he initially assumed the variable to be a nova; but after checking earlier plates and drawing a light curve- brightness versus time- he realized the star displayed Cepheid characteristics and began to calculate its period and absolute brightness. Comparing these calculations with the star's apparent

luminosity using the period-luminosity relation, which related the period of a Cepheid to the star's absolute luminosity. Finally, by comparing the apparent and absolute luminosities of the Cepheid, Hubble estimated the distance to the star and the nebula to be about one million light years. Soon thereafter, Hubble discovered other Cepheids and repeated the process. It was these measures of distance that provided the empirical evidence needed to resolve the debate over the existence of external galaxies:

> Within a year or so, Hubble had accumulated enough evidence from the Cepheids and other distance indicators to convince almost all astronomers that the outer regions of the Andromeda Nebula consisted of clouds of stars and that the nebula was indeed an external galaxy. Following these finds and similar ones for other, large and so presumably nearby spirals, the debate over the existence of visible galaxies was all but at an end.[27]

By examining other stars in these nebulae, Hubble concluded that their upper limit of luminosity was approximately constant. Therefore, the brightest stars in nebulae could be used as the second criterion for determining distance for all nebulae in which any stars could be resolved.[28] In 1928, Hubble began applying this method to nebulae whose velocities had been obtained earlier mostly by Slipher and later by Humason, and discovered that a linear relationship existed between red-shifts and distance. He published his results in a 1929 *Proceedings of the National Academy of Science* article entitled "A Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae."

---

brightness provided Hubble with an estimate of the star's distance, as well as the distance to the nebula to which it belonged.

[27] Smith, 1990, p. 56. Also, for a more detailed account of the "great debate," see Smith, 1982; and Hoskin, 1976.

[28] Hubble, 1929, p. 168-169.

In this short paper, Hubble concluded that, "The results establish a roughly linear relation between velocities and distances among nebulae for which velocities have been previously published, and the relation appears to dominate the distribution of velocities."[29] However, at this stage, Hubble did not comment on the implications of this result. Since the data only reached to about sixty million light-years, he thought it would be premature to make any inferences about the nature of red-shifts:

> New data to be expected in the near future may modify the significance of the present investigation or, if confirmatory, will lead to a solution having many times the weight. For this reason it is thought premature to discuss in detail the obvious consequences of the present results.[30]

In fact, he only mentions theory in his last paragraph, stating that this relation might represent the de Sitter effect, which predicted observed red-shifts as a result of the apparent slowing down of atomic vibrations in distant nebulae. Nevertheless, Hubble maintained that more evidence would be needed before any consequences could be discussed. He also acknowledged that Milton Humason, at Mount Wilson, had initiated a research program to obtain red-shifts for more distant nebulae, out to a limit that could be observed with confidence.[31]

The second phase of investigating the velocity-distance relation, which, in Hubble's view, extended the data out to the limits of the Mount Wilson 100-inch reflector, lasted from 1928 to 1936. During this time, Humason assembled the spectra of

---

29 Hubble, 1929, p. 173.
30 Hubble, 1929, p. 173.
31 Hubble, 1929, p. 173.

nebulae while Hubble estimated their distances. Again, Hubble used previously collected data to correlate other methods for determining distance. One of these criteria was the total luminosities of the nebulae themselves. In effect, this criterion was statistical, but the mean results of a large collection were quite reliable in Hubble's opinion.[32]

Another criterion Hubble used to acquire distances was the brightest nebulae in clusters. This method was particularly valuable in determining the distance of nebulae in which stars could not be detected. Since he considered the clusters of nebulae to be so similar, Hubble believed that the mean luminosity of the ten brightest, or even the individual luminosity of the fifth brightest, of each cluster would offer a precise and convenient measure of distance.[33] Hubble observed clusters that varied from groups of large, bright nebulae to clusters of tiny, extremely faint nebulae. Again, because of their similarity, it appeared as if an observer was looking at a single cluster from different distances. As a result, Hubble believed the relative distances of the nebulae could be estimated by their apparent characteristics, and the actual distance to any one would lead to the actual distances of them all. Hubble thought that this criterion could be applied to obtain the greatest distances, with existing instrumentation, and that the analysis of this collection would indicate the absolute characteristics of the nebulae in general.[34]

By using these criteria, Hubble and Humason were able to obtain data from ever-greater distances. By 1931, Hubble had decided that the data had confirmed the linear relation out to a distance of 150 million light-years.[35] However, Hubble was still reluctant to discuss interpretations of the relation at that time. In his 1931 paper, "The Velocity-

---

32 Hubble, 1937a, p. 10.
33 Hubble, 1937a, p. 11.
34 Ibid.

Distance Relation Among Extra-Galactic Nebulae," Hubble again stated that he desired to emphasize the observational results, and was constrained to "describe the 'apparent velocity-displacements' without venturing on the interpretation and its cosmological significance."[36] What seemed important to Hubble at this time was that, independent of theory and aside from its cosmological significance, the ratio of red-shift to distance was constant, and that the relation offered a new method of determining the distance of individual objects in which the percentage of errors actually decreased with distance.[37]

---

35 Hubble, 1934a, p.9.
36 Hubble and Humason, 1931, p.80.
37 Hubble and Humason, 1931, p.75-77.

Chapter 2: Cosmological Models and Interpretations of Red-Shifts

The historian of science Norriss Hetherington argues that in the mid-1930's, most cosmological theories were derived from the field equations of general relativity. Prior to 1929, the universe was assumed to be static. The idea of an expanding universe was not a serious consideration among astronomers, and only static solutions to Einstein's equations were investigated.[38] However, as we shall see, these solutions became highly problematic by the end of the 1920s, and, as J. D. North asserts, "the supposition that the universe is, on the whole, static did not survive into the 1930s."[39] Before we discuss the reasons for this shift of viewpoint, let us first consider the dominant cosmological models of the early twentieth century.

In 1917, Einstein and William de Sitter, a Dutch astronomer and director of the Leiden Observatory, derived the first solutions to Einstein's equations of general relativity based on the assumptions that the universe was homogeneous, isotropic and static. The universal models of Einstein and de Sitter were termed System A and System B respectfully. Einstein's model contained matter and displayed no systematic motion. Hetherington maintains that at the time it was conceived there was some observational support for this model. However, there was also some observational support for de Sitter's model, which could not contain matter and remain stable. The model would also have an apparent velocity-distance relationship due to diminishing frequency of light vibrations in relation to the increasing distance of nebulae.[40] Slipher's discovery of the

---

38 Hetherington, 1982, p. 43. However, as we shall see, Alexandre Friedmann did produce non-static solutions in 1922 and 1924, but they went unnoticed until the end of the decade.

39 North, 1990a, p. 92.

40 Hetherington, 1982, p. 43. See also footnote 18.

high radial velocities of nebulae lent support to the possibility that a relation between velocity and distance existed, and thus to the possibility of de Sitter's model.

By 1930, however, it became clear to astronomers that neither System A nor System B provided an accurate representation of the actual universe when compared with the observational evidence. Since Einstein's model (System A) displayed no systematic motion, it was completely unable to account for red-shifts and therefore became unacceptable after Hubble demonstrated the existence of a velocity-distance relation in 1929. Since De Sitter's model (System B) could only be applied to a universe in which the mean density of matter was close to zero, it was regarded as unacceptable when the density of matter was estimated to be too high for the model to be a good approximation of the universe.[41] In other words, Einstein's model was abandoned because it could not account for red-shifts, and de Sitter's because it neglected the existence of matter.[42] Since these were the only solutions given serious consideration at the time and it became evident that neither was possible, this presented a dilemma in theoretical cosmology.

The dilemma resulted from only considering these two specific models, and, as Arthur Eddington saw it, from only considering static solutions. At a meeting of the Royal Astronomical Society on 10 January 1930, Eddington commented on the dilemma and supposed, "the trouble is that people look for static solutions."[43] Once astronomers decided the static solutions were untenable, they were forced to look for alternative systems. In Hetherington's opinion, had they been doing so earlier, they may have noticed Alexandre Friedmann's paper from 1922 regarding the curvature of space, or

---

41 Ibid., p. 43-44. Also see Smith, 1982, p. 186.
42 Hubble, 1936b, p. 198.
43 Hetherington, 1982, p. 44.

Howard Robertson's paper from 1929 dealing with red-shifts of spiral nebulae due to recessional velocity in Euclidean space.[44] However, these papers went unnoticed, as did Georges Lemaître's 1927 paper on a homogeneous universe of constant mass and increasing radius that accounted for the radial velocity of nebulae.[45] Once the Einstein and de Sitter models were judged to be inadequate, these earlier works were rediscovered; and, especially the latter, were viewed by some, including Eddington and de Sitter, as a solution to the theoretical dilemma.[46]

Helge Kragh contends that a paradigmatic change in cosmology resulted from the "discovery" of Lemaître's work in 1930-31. This may have very well been the case, since, as Kragh points out, from this point on models that did not incorporate the new idea of universal expansion became obsolete.[47] However, after the change of view, there was still resistance among some astronomers. North maintains, "So strong was this conviction that we inhabit a static universe that long after the discovery of the galactic

---

44 Hetherington, 1982, p. 44-45. Smith and Ellis also make similar contentions. Smith says, "While non-static solutions to the field equations made no 'sudden and dramatic appearance' in cosmology, it was not until the late 1920's that there seemed to be any pressing reason for introducing such a solution. In consequence, solutions such as Friedmann's of 1922…had provoked little interest when they were first expounded." Smith, 1982, p. 188. Similarly, Ellis states, "In the period from 1917 to 1930, everyone knew the universe was static. Consequently when the idea of the evolving universe was proposed, first by Friedmann and then by Lemaître, it was ignored or resisted: the scientific community was not ready for this concept." Ellis, 1990, p. 101.

45 The historian Helge Kragh would, I think, agree with Hetherington as to why these early works dealing with non-static solutions were overlooked, but Kragh develops the point a little further. Not only were astronomers not actively looking for non-static solutions prior to 1930, but the works of men such as Weyl, Lanczos, and Friedmann were considered purely mathematical solutions and not given any particular physical significance. Lemaître's work was not only seen as a possible non-static solution because Lemaître reminded Eddington of his earlier work at a time when expanding models were becoming actively investigated, but Kragh argues there was a defining characteristic behind his work that distinguished it from the other non-static solutions. This distinguishing characteristic was the fact that Lemaître's work "was a serious attempt to give a physically realistic cosmology, not only a mathematical exercise in general relativity." (Kragh, p. 124) Unlike Friedmann, Lemaître explicitly connected his theory with the recession of the nebulae and was prepared to relate the mathematics to astronomical observations. See Kragh, 1987, p. 120-125.

46 See Smith, 1982, p. 186-187.

47 Kragh, 1987, p. 127.

red-shifts there was a veritable industry in finding alternatives to interpreting them as Doppler shifts indicative of real velocity."[48]

As Hubble recalled in 1936, possible universal models involved two undetermined quantities, the cosmological constant and the radius of curvature of space. By assigning different values to these quantities, the theorist would obtain various classes of possible universes, of which the actual universe would be included. The observer could then determine the actual values of the quantities and narrow the range of possible universes even further, to those that were nonstatic with a radius of curvature that varied with time. The equation itself did not indicate whether the actual universe would be expanding or contracting, but the observed red-shifts were generally accepted as direct evidence that the universe was expanding.[49]

Although some astronomers recognized Lemaître's model as a solution to the dilemma, it did not gain universal acceptance. As Hetherington points out, one of the men in opposition to Lemaître's model was the British cosmologist, E.A. Milne. Milne's major objections to the theory dealt with the notion of spatial curvature and the lack of reasoning in the theory for predicting the universe's expansion rather than contraction, since, in Milne's opinion, both were equally probable based on relativity theory.[50] Milne wanted to know why the universe was expanding and why the nebulae had velocities proportional to their distances. As an alternative solution, he proposed a purely kinematical model in which an initial group of nebulae were randomly moving about in

---

48 North, 1990b, p. 13. Also see Ellis, 1990, p. 105.
49 Hubble, 1936b, p. 199.
50 Hetherington, 1982, p. 45.

Euclidean space. The nebulae with the highest velocities would then have receded further, giving the appearance of an expanding group with a velocity-distance relation.

Another proposed explanation under consideration was that of Fritz Zwicky, a physicist at the California Institute of Technology. Zwicky did not believe that the universe was expanding at all. He maintained that the observed red-shifts could be the result of interactions between light quanta and the matter in space, instead of indicating actual recession.[51] Following Zwicky's theory, the velocity-distance relation was explained by a new, unknown principle of physics, perhaps similar to a gravitational drag on light that increasingly affected more distant nebulae in a manner that could produce the observed red-shifts.

Hubble's Evaluation of the Significance of Red-Shifts:

Once the velocity-distance relation was established, Hubble began to focus on its significance to cosmology, which depended on the interpretation of red-shifts. His stated goal was to determine which cosmological model was the most compatible with observations. The effects of red-shifts could be described in several ways. For instance, the light from distant nebulae could be redder than normal, the light waves were longer, the vibrations were slower, or the light quanta could lose energy. However, as Hubble stated in 1934,

> Many ways of producing such effects are known, but of them all, only one will produce large red-shifts without introducing other effects which should be conspicuous but actually are not found. This one known permissible explanation interprets red-shifts as due to actual motion away from

> the observer. Rapid motion of recession drags out the light waves and, as it were, lowers the pitch. We may say with confidence that red-shifts are due either to actual motion or some hitherto unrecognized principle of physics. Theoretical investigators almost universally accept the red-shifts as indicating motion of recession of the nebulae, and they are fully justified in their position until evidence to the contrary is forthcoming.[52]

Following the interpretation of actual motion, the nebulae are receding from us with a velocity proportional to their distance. Based on the velocity-distance relation, Hubble found that the velocities increased by roughly 100 miles per second for each million light-years of distance.[53] Following this recession backwards in time led to the assumption that the nebulae were once packed together in our particular region of space, and at a particular instant they started rushing away in all directions at various speeds. According to the data available in the early 1930s, Hubble estimated that this took place about two billion years ago. This short time scale presented one of the first problems with the interpretation of red-shifts as velocity shifts. At the time, it was considered a small fraction of the estimated age of some stars, and therefore presented a dilemma.[54]

---

51 Hetherington, 1982, p. 46.

52 Hubble, 1934a, p. 13.

53 Ibid. The value of the increase in nebular velocity is dependent on the value assigned to the Hubble constant.

54 Kragh argues that from the start it was evident that the idea of the expanding universe posed as many problems as it solved. The relatively young age of the universe and the beginning of expansion were two such problems. There were other ways to interpret the time frame of expansion that could compensate for the young universe. For instance, if the universe existed before the expansion began, then the age would not be as much of a problem. The Lemaître-Eddington model in 1930 provided such an interpretation. Their model contained an expanding universe that evolved gradually from an already existing pre-universe and therefore had no proper beginning. Therefore, the time-scale of expansion would be nothing compared to the age of the universe itself. However, as Kragh argues, the men soon came to disagree on the aspect of creation. For Eddington the idea of a beginning of Nature was philosophically repugnant, but by 1931 Lemaître introduced the notion in a realist sense. Lemaître argued, by means of quantum theory, that the universe came about by the division of a primeval atom with a mass of the entire universe. The notions of space and time would also take on meaning at the same time. Based on the idea of the disintegrating primeval atom, Lemaître did not believe the time-scale dilemma had to arise. Kragh, 1987, p. 127-132.

Hubble believed that the solution to this dilemma would have to await further observations. Since nebulae could be observed and photographed at distances where their spectra could not be recorded, Hubble saw the need for more observations with larger telescopes. Hubble believed that this would make it possible to test the velocity-distance relation further, as well as the interpretation of red-shifts as velocity-shifts. Therefore, Hubble refrained from committing himself to the interpretation and used the convenient term "apparent velocity" to describe their effects.[55] In his 1934 paper, *Red-Shifts in the Spectra of Nebulae*, Hubble then began to look at the behavior of red-shifts.

Regardless of their interpretation, Hubble assumed red-shifts to behave as velocity-shifts, that is, that the displacements of lines, expressed as fractions of their normal wave-length, are constant throughout the whole of a given spectrum. He could then measure red-shifts by the ratio $d\lambda/\lambda$, where λ is the wave-length of the unshifted line. Next, Hubble calculated the effects of red-shifts on the bolometric luminosity, the total radiation measured as if the observer was outside the Earth's atmosphere.

North explains the dilemma as follows: "If it is possible to speak of the age of the universe at all, it will be agreed that this cannot be less than the age of any constituent part, or the duration of any single constituent process. Now when the reciprocal of the 'constant' in Hubble's Law was first compared with estimates of the age of the Earth- then, as now, taken to be of the order 10^9 or 10^{10} years- the similarity was striking. Unfortunately there was a firm belief that Jeans had established the age of the Galaxy at 10^{12} or 10^{13} years, basing his arguments on the dynamics of star systems. Here was a paradox under which the evolutionary theories laboured for many years. How can the Galaxy be older than the universe itself?" North, 1990a, p. 224. Actually, the dilemma was resolved years later when it was realized that Hubble had made an error in his method of determining distances by the use of Cepheid variables. In 1952, the American astronomer, Walter Baade, discovered that there are two types of Cepheids with different period-luminosity relations. This discovery did not affect the measures of distance within our own Galaxy, because they were determined using Type II Cepheids. However, the Cepheids observed in other galaxies are of Type I, and their absolute luminosity was underestimated. This also meant that their distances were underestimated. Sciama explains the importance of this discovery: "The change has been a substantial one. Distances derived by Hubble should be multiplied by about 5, thus greatly increasing the scale of the Universe…Hubble's constant is now believed to be about 10 billion years. This is greater than the accepted ages of the Earth and Sun, and comparable with the ages of the oldest star clusters. Thus there is no longer any difficulty in supposing that the Universe was once very dense." Sciama, 1971, p. 41-46.

[55] Hubble, 1934a, p. 14.

Radiation from a nebula can be described in terms of light quanta, or tiny parcels of energy that stream out in all directions. Apparent luminosity can then be measured by the rate at which the quanta reach the observer. All quanta have the characteristic such that, energy x wave-length = constant, therefore a red-shift, regardless of interpretation, decreases the energy of a quanta by increasing the wave-length. The amount of energy decrease is constant throughout the spectrum and can be represented by the factor $(1+d\lambda/\lambda)$. This effect of red-shifts was called the energy effect.[56]

Another effect of red-shifts Hubble calculated on the bolometric luminosity was termed the number effect. Again, if red-shifts were assumed to represent actual motion, then the rate at which quanta reach an observer would be less than the rate at which they left a nebula. The path that they travel would be lengthened and the rate would be reduced by the factor $(1+v/c)$, where v is the velocity of the nebula and c is the velocity of light. Since the rate that the quanta reach the observer is reduced, the apparent luminosity is also reduced. The factor was known as the number effect, and operates only in the case of actual motion. In expanding models of the universe obeying the relativistic law of gravitation, the ratio v/c was replaced by $d\lambda/\lambda$, and, like the energy effect, the factor became equal to $(1+d\lambda/\lambda)$.[57]

Next, Hubble calculated the effects on photographic magnitude. In order to actually apply the results of the observed data, Hubble found it necessary to trace the effects through the Earth's atmosphere and down to the range in the spectrum that is recorded on a photographic plate. The reason for this was that the atmosphere had

---

56 Hubble, 1934a, p.14-15.
57 Hubble, 1936a, p. 535.

selective absorption, the telescope had selective reflection, and the photographic plate had selective sensitivity.[58] The number effect was the same for photographic and bolometric luminosity, because it was non-selective, or the same for all wave-lengths. It merely reduced the vertical, or intensity, scale of the energy-distribution curves. However, the energy effect was selective, the effect at each wave-length depended on the wave-length itself, and it changed the horizontal or wave-length scale of the energy-distribution curve. Therefore, he introduced a K term into the expression of the energy effect to represent the reduction from bolometric to photographic magnitude.[59]

To conclude his investigation in his 1934 paper, Hubble offered some possible interpretations of his calculations. He summarized them as follows:

> If red-shifts increase linearly, and the distribution of nebulae is uniform, then *either* the K term is negligible, *or* the red-shifts do not represent motion. Again, if the red-shifts represent motion, and the corrections are properly calculated, then *either* the rate of increase of red-shifts diminishes with distance, *or* the density of nebular distribution increases with distance…Aside from these possibilities, about the only definite conclusion that can be drawn from the investigation is that the rate of increase of red-shifts probably does not accelerate with distance.[60]

At this point, Hubble refrained from choosing a decisive interpretation. He reckoned it doubtful that a solution could be obtained with the instruments that were available, but he was optimistic that data acquired with the 200-inch telescope, then

---

58 Ibid.
59 Ibid.
60 Hubble, 1934a, p. 17.

under construction, would furnish the critical tests of the interpretations and extend the limits of the velocity-distance relation.[61]

In 1935, Hubble published the paper entitled, "Two Methods of Investigating the Nature of the Nebular Red-Shift" with Richard Tolman, a theoretical physicist from the California Institute of Technology and the author of *Relativity, Thermodynamics and Cosmology*. By 1934, Tolman had begun developing the mathematical foundations of relativistic cosmology, and calculating certain features of the homogeneous, expanding universe, such as a formula that expressed the relative numbers of nebulae that should be observed in surveys when apparent magnitudes are reduced to a single epoch.[62]

In their joint paper, the authors began by identifying the most obvious interpretation of red-shifts. This interpretation regarded red-shifts as being directly correlated with a recessional motion of the nebulae, and was commonly adopted in the treatments of nebular motion that were made with the help of the relativistic theory of gravitation and in the purely kinematical treatment proposed by E. A. Milne.[63] However, they did not neglect the possibility that the red-shifts could be attributed to some other cause, despite the fact that, in their view, no other satisfactory account had been proposed. Therefore, both authors wished to keep an open mind on the matter until further evidence became available, and continued to use the term "apparent velocity" when presenting observational findings. At the same time, however, both maintained the

---

61 Ibid.
62 Hubble, 1936b, 199-200.
63 Hubble and Tolman, 1935, p. 302-303.

opinion that if red-shifts were not due to recessional velocity, they must be explained by some entirely new principle of physics.[64]

The main purpose of this paper was to study in detail the procedure that had to be adopted in applying two theoretical methods of investigating the nature of red-shifts to observations. The first method depended on the relation of nebular dimensions and observed luminosities, and the second depended on nebular counts and observed luminosities, both relations were to be expected for a homogeneous distribution of nebulae under the alternative assumptions of mutual recession or of some other cause of red-shifts.[65]

The authors divided the paper into three parts. Part one dealt with the relations that were to be expected to hold in the case of actual recession. For this case, the authors used a homogeneous expanding model, which obeyed the relativistic laws of gravitation, to represent the actual universe. The reasons Hubble and Tolman gave for selecting this model were, the principles of relativity furnished the most adequate description of gravitational phenomena at the time, and they were definite and provided a sufficient basis for calculating relations needed for the proposed tests of the nature of red-shifts.[66]

Part two of the paper dealt with similar relations, except in this case the red-shifts were assumed to be caused by some factor other than recession. Just as the cause of red-shifts in this case was unknown, the authors also expressed a degree of uncertainty about

---

64 Hubble and Tolman, 1935, p. 303.

65 Ibid. Hetherington maintains that neither method had been previously developed in detail, and that Hubble and Tolman attempted this approach with the goal of discriminating between possible models of the universe, as well as linking theory and observation. Hetherington, 1982, p. 48-50.

66 Hubble and Tolman, 1935, p. 304.

which model should be used to represent the actual universe.[67] For definiteness, Hubble and Tolman used a static Einstein model, combined with the assumption that photons lose energy by some unknown effect as they travel to the observer. Although the static Einstein model was known to be unstable, the authors stated that they adapted the model for their interests to investigate the possibility that red-shifts may not be caused by recession.[68] In this case, the Einstein model was used as a limiting case of models in which the effect of expansion or contraction was negligible compared to the actual, unknown cause of red-shifts. The third part of the paper contained a preliminary statement about the status of observational work taking place at the time.

In the case of the non-recessional theory, the authors found that approximate agreement with the data could be attained provided they disregard the possibility of any appreciable light-absorption on the way from the nebulae, and at the same time neglect any possible effects of spatial curvature. In the case of the recessional theory, which predicted larger effects from the red-shifts, they found that approximate agreement with the data would seem attainable only by making use of possible effects from spatial curvature.[69]

In regards to the nature of spatial curvature needed in the recessional theory, the authors noted that the effects would be in the right direction for agreement only for the case of a closed model with real spatial curvature. This result was of interest in connection with Milne's theory, because he had shown that his kinematical model would

67 Hubble and Tolman, 1935, p. 304.
68 Hubble and Tolman, 1935, p. 304.
69 Hubble and Tolman, 1935, p. 334.

work with negative curvature.[70] Hubble and Tolman argued, however, that the discrepancies could be reduced if a positive value was assigned, while a negative value increased them. As Hubble recalled in a later paper from 1936, they therefore ruled out negative curvature, indicating an open universe, and restricted the types of possible universes to those with positive curvature.[71] If the red-shifts were velocity-shifts, then they believed the universe was closed and had a finite volume and contents.

The amount of spatial curvature required in a homogeneous cosmological model seemed to lead to another dilemma. The amount would require an averaged-out density in the universe that was several thousand times as large as that which could be computed from the masses usually assigned to the nebulae alone. A density that high would necessitate either large amounts of unknown intergalactic matter in a form that would not lead to absorption, or very large masses of the nebulae themselves.[72]

Hubble and Tolman argued that another problem with the recessional theory was in regards to the assignment of constant properties to the nebulae over long periods of time. If the recessional theory was adopted, the period of time required for light from distant nebulae to reach the observer would be nearly comparable with the time of cosmic expansion, estimated by the authors to be on the order of 10^9 to 10^{10} years.[73] This would require a relatively fast evolution time for nebulae, and therefore the assignment of constant properties did not seem justified to them. On the other hand, with the non-recessional theory, there was nothing to indicate a short time scale for nebular evolution, which may have seemed to be an advantage for that theory.

---

70 Ibid., p. 335. See also Hetherington, 1982, p. 49.
71 Hubble, 1936b, p. 200.
72 Hubble and Tolman, 1935, p. 335.

The authors summarized their conclusions as follows:

> Our conclusions as to the present status of nebular counts can therefore be summarized by stating that it might be possible to explain the results on the basis of either a static homogeneous model with some unknown cause for the red-shift or an expanding homogeneous model with the introduction of effects from spatial curvature which seem unexpectedly large but may not be impossible. This statement must be made very tentatively, however, and may need serious modification in the light of further counts, of new auxiliary data for treating the counts, or of better analysis of the present data.[74]

Hubble and Tolman's paper did not result in a choice between alternative models, but rather a plea for additional data.[75] However, although they didn't achieve their goal, their efforts were not entirely in vain, as Hetherington points out:

> The upshot of the investigation by Hubble and Tolman was that they could not choose between nonrecessional and recessional models. In the case of the recessional theory though, the appropriate relativistic model did not seem to be the one similar in observational properties to Milne's kinematic model, because the observed increase of nebular counts with distance indicated a positive curvature of space…Their 1935 article thus formulated methods for interpreting the nature of the red-shifts but it did not arrive at definite conclusions concerning the nature of the red-shifts and the corresponding model of the universe.[76]

---

73 Ibid., p. 336.

74 Hubble and Tolman, 1935, p. 335.

75 Hetherington contends that this was a result of their method being impractical at the time. "Perhaps the method would be applicable following improvements in techniques used with the 100-inch telescope or upon completion of the new 200-inch telescope, but it was not practical in 1935." Hetherington, 1982, p. 50.

76 Hetherington, 1982, p. 51.

Although Hubble and Tolman did not reach any definite solutions in their paper, it did not stop Hubble from pressing on. In 1936 Hubble followed another approach to evaluating the interpretations of red-shifts, which involved calculating the effects of red-shifts on the distribution of nebulae. Hubble judged that surveys show the distribution of nebulae over the sky to be moderately uniform, except as affected by local obscuration within the galactic system and by the occasional great clusters. The distribution in depth also appeared to be approximately uniform when tentative corrections were applied for the effects of red-shifts on the apparent luminosity of the nebulae. However, Hubble believed this latter conclusion was less reliable and required further investigation before it could be used with confidence in cosmological theory.[77]

Hubble argued that the observational results were consistent with the assumption that the apparent departures from large-scale uniformity, Δm, were linear functions of distance and represented the effects of red-shifts on apparent luminosity. As a result, in Hubble's opinion, a comparison of the observed value of B in the relation $\Delta m = B d\lambda/\lambda$ with values calculated on various interpretations of red-shifts should indicate which interpretation was more probable.[78]

By comparing observations with theory, Hubble wrote that the observed result, $B = 2.94$, was accounted for if red-shifts were not velocity-shifts.[79] This interpretation was consistent with the data whether the nebulae were pictured as scattered through an infinite universe or if a homogeneous model obeying relativistic laws of gravitation was adopted. In this latter case, the expansion and spatial curvature were assumed to be

---

77 Hubble, 1936a, p. 517-518.
78 Ibid., p.534.
79 Ibid., p. 542.

negligible or zero. The interpretation also assumed that the large-scale distribution was uniform, that the nebulae maintained fairly constant luminosities over the time interval involved in the observations, and that red-shifts were an approximately linear function of distance.[80]

The first assumption, real uniformity, followed from the small values of the observed departures from apparent uniformity. The departures were no greater than the effects of red-shifts. When the observations were corrected for the effects of red-shifts, the result was uniformity. Further corrections indicated that the real distribution density increased outward, and this was unacceptable to Hubble because it implied a unique position for the observer. This was the reason that he postulated uniform distribution and assumed space absorption to be negligible. The observed departures could then be attributed to the effects of red-shifts alone.[81]

Hubble's assumption that the luminosities of nebulae remained constant over the time interval stemmed from the question of whether or not luminosities of nebulae changed during the time it took for light to travel from the limit of the deepest surveys to the limit of the shallowest surveys. This time interval was calculated by Hubble to be about 250 million years. Hubble maintained, "no definite information was available…Nevertheless, very few students will hesitate to adopt the assumption that systematic variation in so short an interval will be inappreciable. Nebulae are enormous systems, and it is reasonable to suppose that their evolution is correspondingly slow."[82]

---

80 Hubble, 1936a, p. 542.
81 Ibid., p. 542-543.
82 Ibid., p. 543.

Hubble's third assumption, that the velocity-distance relation is sensibly linear, resulted from the close representation of the data with the expression $\Delta m = 3d\lambda/\lambda$, that would be used to represent corrections for the effects of red-shifts on apparent luminosity under the assumption that red-shifts are not velocity-shifts. When red-shifts were assumed to be velocity shifts, the larger correction, $4d\lambda/\lambda$, had to be used. This latter value introduced departures from linearity in the velocity-distance relation. Hubble judged that the data were sufficient to show the existence of such departures, but numerical values were still uncertain. Hubble emphasized, however, that the empirical value was precisely the effect of recession that was introduced as an additional assumption not necessarily demanded by the observations themselves.[83]

Hubble then compared the observational results with homogeneous models of the universe that obeyed the relativistic laws of gravitation. The principles and methods he used were those discussed in his earlier paper written with Tolman. The data were compared for two cases. The first assumed red-shifts were not velocity-shifts, and the second assumed they were velocity-shifts and that they disclosed the rate of expansion of the universe.

Again, the model Hubble used in the first case was Einstein's static universe in which the rate of expansion was considered inappreciable. When compared with this model, the observations were thoroughly consistent, provided that spatial curvature was negligible. On the other hand, the expanding models associated with the second case

---

[83] Hubble, 1936a, p. 546.

were definitely inconsistent with the observations when curvature was assumed to be negligible.[84]

If the discrepancies between observations and expanding models were due to curvature, then, Hubble argued, the effects of curvature had to be just sufficient to balance the number effects that were introduced by the assumption of actual motion. The curvature then had to be real, positive, and very large. This positive curvature required the universe to be closed, finite and small. The radius of curvature would then be just less than the estimated distance to the faintest nebulae recorded, which indicated that a large fraction of the entire universe had been observed. A radius of this order would also require a great increase in the density of the universe. Therefore, the validity of this model, as representing the actual universe, may have depended on the question of density.[85]

Hubble considered that the high density implied by a model of those dimensions presented a problem for the theory. If the estimations of density were reliable, then the evidence would rule out the necessary radius of curvature. However, as Hubble pointed out, the crucial data were surrounded with uncertainties, and other forms of matter might exist between the nebulae.[86] Since the galactic system, as compared with other spirals, faded outward to undefined boundaries, it was reasonable to suppose that the mean density of matter in internebular space was thousands of times less than the local density.[87] It could at least be assumed that the density was greater than zero. However, the observations indicated that space absorption was negligible. This condition ruled out

---

84 Ibid., p. 548.
85 Hubble, 1936a, p. 551-552.
86 Ibid., p. 201.

many forms of matter, for example, dust, meteoritic material, and highly ionized gas; but Hubble believed certain forms could still be possible.[88]

After this investigation was completed, Hubble asserted that the observations could be fitted into either of two quite different types of universes. If red-shifts were velocity-shifts, then the model was closed, small and dense. It would be rapidly expanding, but over a long period the rate of expansion would decrease. A large fraction of the universe would have been observed with the instruments in existence at the time.[89]

Alternatively, if red-shifts were not velocity-shifts, the observable region became insignificant. The velocity-distance relation remained linear, the distribution of nebulae was uniform, there was no evidence of expansion, no trace of curvature, and no restriction of the time scale; but the sample seemed too small to indicate the actual universe we inhabit.[90]

Hence, the investigations out to limits of the 100-inch telescope presented a choice between a curiously small-scale universe and a thereunto unrecognizable principle of nature. Again, Hubble contended that a definitive choice, based on observations, would not be possible until results with the 200-inch reflector became available.[91] In the meantime, Hubble emphasized certain features about the two solutions. One of these features was the fact that, in Hubble's opinion, no other satisfactory interpretation of red-shifts had been formulated, other than that derived from the familiar Doppler effects that assumed that red-shifts measured the rate of expansion.

---

87 Ibid., p. 552.
88 Ibid.
89 Hubble, 1936a, p. 553.
90 Ibid.
91 Ibid.

However, most of the discrepancies and unexpected features of the investigation were introduced by the additional assumption of recession. For instance, the velocity-distance relation deviated from linearity by the exact amount of the postulated recession. Also, the distribution of nebulae departed from uniformity by the exact amount of the recession, and, in order to compensate for these departures, spatial curvature was introduced that was also the exact equivalent of the recession. In other words, the discrepancies could all be attributed to the introduction of the recession factor in one form or another. Hubble believed that " unless the coincidences are evidence of an underlying, necessary relation between the various factors, they detract materially from the plausibility of the interpretation."[92] Although the curvature could compensate for these discrepancies with observations, it also introduced other problems with the model, such as a very large density and a small scale of volume and time. Hubble supposed these remarkable features would require decisive evidence to gain acceptance for the model.

Observing the number of nebulae in surveys to successive limits of apparent magnitude indicated an approximately uniform distribution. However, the effects of red-shifts on apparent magnitudes also effected the apparent distribution. By assuming a precisely uniform distribution, and that departures were represented by the combined effects of red-shifts and observational errors, Hubble evaluated the calculated corrections based on alternative interpretations as a way of testing the interpretations themselves. Since the numerical results were different, Hubble used the observed departures to try to identify the correct interpretation.[93]

---

92 Hubble, 1936a, p. 554.
93 Hubble, 1936b, p. 193.

As mentioned earlier, regardless of their interpretation, red-shifts affected apparent luminosities by the amount known as the energy effect. When red-shifts are interpreted as velocity-shifts, an additional correction, the number effect, also had to be applied. Therefore, Hubble reckoned that the number effect furnished a critical test of the interpretation. Since the effects of red-shifts on magnitude, Δm, was a linear function of distance, and red-shifts were a linear function of distance, the two could be expressed as a linear function of each other. The calculated values, $\Delta m = 4\ d\lambda/\lambda$ (for recession) and $\Delta m = 3\ d\lambda/\lambda$ (no recession), could then be compared with the observed value of Δm, which Hubble estimated was equal to $2.7\ d\lambda/\lambda$.[94]

Although the observed value was less than both of the calculated values, the data were in closer agreement with the relation based on the interpretation of no motion. After careful examination of possible sources of errors, Hubble concluded that the observations could probably be accounted for if red-shifts were not velocity-shifts, and that if they were velocity-shifts then some vital factors must have been neglected.[95]

Consequently, if red-shifts were velocity-shifts, then the universe would be very small and packed with matter. If red-shifts were not velocity-shifts, then these discrepancies would disappear, but the cause of red-shifts and its implications would be unknown. Hubble was faced with a choice between a suspicious model of the universe and an unrecognizable principle of physics. Once again, because of uncertainties, he refrained from making a definite decision about the true nature of the universe.

---

94 Ibid., p. 167-197.
95 Hubble, 1936b, p. 197.

Chapter 3: Bringing the Pieces Together

As we saw in the last chapter, in the latter part of the 1920's and early 1930's, Hubble went to great lengths to establish his velocity-distance relation. In doing so, he carefully avoided discussions of theory and chose to emphasize empirical data. Initial investigations revealed an approximately linear relation. In order to establish the true nature of the relation Hubble wanted to extend his investigations out to still greater distances. The linear relation survived the successive tests and was found to hold, out as far as observations could be made with existing telescopes. Therefore, the investigations led to a new law, eventually credited to Hubble, which stated that there is an approximately linear relation between red-shifts and distance.

The historian of science, Robert Smith, maintains that there were four main reasons for the acceptance of Hubble's linear velocity-distance relation.[96] First, numerous astronomers had already attempted to establish a relation between red-shifts and distance by observational means prior to 1929, so Hubble was not breaking any new ground. However, in his colleagues' opinions, his work was an improvement on these earlier attempts because he determined distances in ways that were regarded as providing the most accurate measures. Secondly, Hubble's reputation as an authority in extra-galactic astronomy and the fact that he utilized the world's most powerful telescope, provided additional credibility for his nebular research and the accuracy of his velocity-distance relation. Thirdly, de Sitter, who also derived a linear relationship from the observations available to him in 1930, quickly confirmed Hubble's relation. Smith contends that de Sitter's reputation as one of the most eminent and influential astronomers "added

---

[96] Smith, 1982, p. 185-186.

credence to the reality of the linear relation." Finally, because the inadequacies of the static solutions of Einstein and de Sitter had become known to mathematical cosmologist by the late 1920s, the relevance of Hubble's relation for cosmological models was almost immediately recognized.

Once his relation gained general acceptance, Hubble then turned to evaluating the interpretation of the relation, which, by the mid-1930s, had become crucially important to deciding on a cosmological model that represented the actual universe. The two quantities involved in the relation were the observed red-shifts in the spectra of nebulae and apparent luminosities. Since Hubble considered absolute magnitudes to be statistically constant, he then interpreted apparent luminosities as measures of distance. Although red-shifts behaved as velocity-shifts, they could not as easily be referred to as such. Therefore, Hubble was careful not to prematurely neglect the possibility of other interpretations and used the term "apparent velocity." Hubble may have been inclined to interpret red-shifts as velocity-shifts, based on their behavior and what he considered a lack of any other acceptable explanation, but he argued that he was approaching the investigation with an open mind.[97] However, careful examination of his writings and philosophy of science demonstrate that this was not the case, and, I will argue that it is often evident that Hubble did not practice what he preached.

The skeptical or cautious approach exhibited in Hubble's writings was a product of his scientific methodology. Hubble believed that until tests could be secured, a healthy skepticism towards unverified speculations was a universal and necessary characteristic

---

[97] Hubble and Tolman, 1935, p. 303.

of research and investigation.[98] It was largely due to this method that cosmology in the twentieth century became an observational science, transformed from a practice that, for centuries, had been based on speculation and a minimum of observational evidence. With this methodology, modern cosmology became a science based on experiment and observations as opposed to dialectics. As Hetherington claims:

> For centuries, cosmology had consisted of speculation based on a minimum of observational evidence and a maximum of philosophical predilection. In the twentieth century, cosmology was transformed into an observational science, largely through Hubble's work. Theories now were subject to verification or refutation to an extent previously unimaginable…The shift in balance from nearly pure speculation largely unadulterated by observational information to a science of cosmology now containing a creative and dynamic tension and a lively interaction between theory and observation was one major change in the practice of cosmology in the twentieth century.[99]

For Hubble, science dealt only with judgments about which it is possible to obtain universal agreement. This was a proposition that Hubble attributed to Norman Campbell's *What is Science*, written in 1921.[100] Agreement was secured by observation and experiment, which Hubble considered "an impersonal court of appeal to which men of all races and all religions must submit if they wish to survive."[101] The formulation of theories to account for data was an ex post facto procedure that offered plausible interpretations of the data. These theories would again be tested by new data, and their validity would be measured by the verification of their predictions. No theory was sacred,

---

98 Hubble, 1936b, p. 5-6.
99 Hetherington, 1996, p. 149.

and if it failed to meet the tests it had to be modified or abandoned. In one of his final papers, written in 1954, Hubble summarized the methodology as follows:

> The subject matter consists of judgments on which universal agreement can be obtained by observation and experiment. The method consists of the interpretation of available data by working hypotheses whose implications are tested by further experimentation. Throughout the process, science is concerned with knowledge that can be freely transmitted, knowledge that comes from observation and experiment.[102]

Applying this method to the science of cosmology, Hubble designated specific, yet complementary, roles for the theorists and observers. By working together and subjecting theories to observational tests, scientists could achieve the common goal of understanding the universe. In "The Problem of the Expanding Universe," Hubble explains this process more clearly:

> Mathematicians deal with possible worlds, with an infinite number of logically consistent systems. Observers explore the one particular world we inhabit. Between the two stands the theorist. He studies possible worlds but only those which are compatible with the information furnished by observers. In other words, theory attempts to segregate the minimum number of possible worlds which must include the actual world we inhabit. Then the observer, with new factual information, attempts to reduce the list still further. And so it goes, observation and theory advancing together toward the common goal of science, knowledge of the structure and behavior of the physical universe.[103]

---

100 Hubble, 1936b, p. 1. Here Hubble quotes from page 27 of Norman Campbell's *What is Science?* (London, 1921).

101 Hubble, 1954, p. 40.

102 Hubble, 1954, p. 41.

103 Hubble, 1942, p. 123.

Underpinning Hubble's scientific methodology was the philosophy of science he expressed in his writings. Hubble believed that the world in which we live was divided into two aspects, the realm of science and the world of values. The realm of science was the public domain of positive knowledge, while the world of values was a private domain of personal wisdom. Knowledge was obtained through observation and experiment, and was also verified by these means. However, wisdom and the meaning of such things as good and evil, beauty and success, were acquired through personal experience. In the world of values there was not an independent authority or method by which universal agreement could be obtained. Therefore, there was not a calculus of values or means of accumulation of wisdom throughout the ages. On the other hand, Hubble believed knowledge was cumulative over time and he quotes the Harvard historian of science George Sarton on the point.

> The saints of today are not necessarily more saintly than those of a thousand years ago; our artists are not necessarily greater than those of early Greece; they are likely to be inferior; and of course, our men of science are not necessarily more intelligent than those of old; yet one thing is certain, their knowledge is at once more extensive and more accurate. The acquisition and systemization of positive knowledge is the only human activity that is truly cumulative and progressive.[104]

Hubble held the realms of science and values to be separate and distinct. In the realm of science, one attempted to explore the universe as it was, or appeared to the

---

[104] George Sarton, *Introduction to the History of Science* (Baltimore, 1927), p. 3-4; quoted in Hubble, 1936b, p. 2.

observer, and was uninfluenced by personal desires or beliefs. The pursuit of pure science was to be characterized by disinterested curiosity.[105]

Although Hubble may have intended to distinguish between the realms of science and values, in his work they tended to merge, and it is often evident that he held theory and philosophical beliefs above observations. It was philosophical beliefs that led Hubble to reject the cosmological model of the universe that was most agreeable to his data, and instead accept a model that was more in line with his fundamental principles. The fundamental principles from which Hubble began his investigations were general relativity and the cosmological principle.

Hubble stated that general relativity predicted a universe that was unstable, either expanding or contracting, and whose space was curved in the vicinity of matter. The amount of curvature depended on the amount of matter. The cosmological principle, which Hubble regarded as a pure assumption, supposed that the universe was homogeneous and isotropic- very much the same everywhere and in all directions.[106] This ensured that there was no favored position in the universe, and that the universe would appear very much the same to all observers, regardless of their location in it.

This postulate was also very important for Hubble's observational approach to cosmology, which consisted of inferring the nature of the universe as a whole from the observed characteristics in the sample of the universe available for inspection. However, sound inferences could be made only if the observable portion of the universe was considered a sufficient representation of the whole, and this could only be accepted by assuming the reality of the cosmological principle.

---

105 Hubble, 1954, p. 38-39.

Once Hubble began to investigate the nature of red-shifts, we have seen that he sought the assistance of theoreticians familiar with the principles of relativistic cosmology, such as Richard Tolman. In 1935 the two coauthored a paper in an attempt to determine which cosmological model best agreed with observations. Their investigation revealed that they could not choose between the recessional and nonrecessional models, however, they did formulate methods for deriving a definite interpretation of red-shifts and the corresponding cosmological model.[107] The reason they gave for not being able to choose between models was the uncertainty of the available data. Although uncertainties were present, the men chose to emphasize these uncertainties rather than discuss which model seemed more consistent. Hetherington argues that this suggests that there was some initial resistance to the implications of the data, which favored a nonrecessional model.[108]

Hubble began with the assumption that red-shifts were velocity-shifts, because he considered it as the only permissible interpretation at the time. Other ways were known that could produce the red-shifts, but these cases would have introduced other effects that were not observed. The red-shifts could be completely explained by stating that light loses energy in proportion to the distance it travels, but the reasons and ways in which the energy was lost would be unknown. Therefore, Hubble could conclude that red-shifts were either velocity-shifts or caused by some unknown principle of physics.[109] However, Hubble judged that the present knowledge of nature was probably very limited.

---

106 Hubble, 1942, p. 124.
107 See footnote 72.
108 Hetherington, 1982, p. 51.
109 Hubble, 1937b, p. 5-6.

Consequently, until absolute proof could be offered for the interpretation, he had to remain open-minded and could not confidently accept the interpretation explicitly.

Initial investigations of the problem led Hubble in circles. If the nebulae were receding, then their light would appear fainter than if they were stationary at the same distance. If distances were known prior to the investigations, then the apparent faintness would reveal whether or not the nebulae were actually receding. However, Hubble's only measure of great distances was the apparent faintness of the nebulae and this was a quantity that was involved in the tests. Therefore, the direct approach to the investigation led Hubble nowhere.

To avoid this problem Hubble chose an indirect approach. This consisted of using the calculated distances and evaluating the data based on two assumptions, one involving the recession of the nebulae and the other holding the nebulae stationary. The results of these investigations would then be compared with respect to their simplicity and internal consistency. The main purpose of this approach, in Hubble's opinion, was to determine the precise formulations of the laws of red-shifts and of nebular distribution, using two different scales of distances corresponding to the alternative interpretations of red-shifts.

When the necessary corrections (energy effects) were applied to the data, it became clear to Hubble that the nonrecessional theory was more consistent with observations. The law of red-shifts remained strictly linear and the distribution of nebulae was found to be uniform and homogeneous, the numbers of nebulae were strictly proportional to the volumes of space and there was no trace of a thinning out.[110] This interpretation of the data, according to Hubble, was direct, economical and very simple;

---

[110] Hubble, 1937b, p. 6-10.

however, it meant that the observable region was an insignificant fraction of the universe, which would be so vast that it could not be described in its totality.[111] Following this interpretation, the universe could be approximated by the Einstein model or a homogeneous expanding model in which the curvature and rate of expansion were inappreciable.

When Hubble analyzed the data under the interpretation of velocity-shifts, it proved to be inconsistent with observations and presented several dilemmas. Since the interpretation involved the recession of the nebulae, the data had to be corrected even further to compensate for the recession factor (number effects). Since the amount of correction increased with distance, the result was a departure from the previously linear law of red-shifts. Hubble observed red-shifts at a given distance to be greater than those predicted by the linear law, indicating that the rate of expansion accelerated with distance. In other words, the rate of expansion had slowed down during the time covered by the surveys. This presented the first dilemma, because it indicated that the age of the universe was much less than had been previously supposed.

When the data had been corrected for the recession factor and the surveys were reduced to a single epoch, Hubble found that the distribution of nebulae was not uniform. The density increased outward in all directions. The observable region appeared to be non-homogeneous, indicating a favored position for the observer. This presented another dilemma for Hubble, to which he had two objections. First, the special position of the observer was an unwelcome and unacceptable situation, because it violated the cosmological principle. Second, two of Hubble's fundamental propositions, the

---

[111] Hubble, 1936a, p. 553.

correctness of general relativity and the cosmological principle, predicted universes that were homogeneous and unstable. A non-homogeneous observable region presented a discrepancy with Hubble's theoretical universe, and therefore could not be a real possibility. To overcome this discrepancy, Hubble assumed that the departures from uniformity were not real, and began to search for neglected factors to compensate for the apparent departures.[112]

We have seen that when Hubble examined the nebular distribution under the energy corrections alone, the data exactly compensated for the apparent thinning out of nebulae at increasing distances, and preserved the uniform distribution. Once the recession factor was introduced, Hubble was confronted with discrepancies. To compensate for the discrepancies Hubble introduced spatial curvature. What Hubble proposed to do was to postulate expansion, then postulate spatial curvature in order to compensate for the departures from homogeneity. Hubble acknowledged that this procedure seemed artificial, but it offered the only means of preserving the theory of a homogeneous, expanding universe.[113]

> To the observer the procedure seems artificial. He has counted the nebulae to various limits, applied only the corrections that are necessarily required (energy-corrections), and derived the quite plausible result of uniform distribution. Now, in testing the relativistic theory, he introduces a new postulate, namely, spatial curvature, in order to compensate the discrepancies introduced by the first. The accumulation of assumptions is uneconomical, and the justification must be sought

---

112 Hubble, 1937b, p. 11.
113 Hubble, 1937b, p. 11.

> in the general background of knowledge. The outstanding argument is the fact that velocity-shifts remain the only permissible interpretation of red-shifts that is known at the present time.[114]

The assumptions and compensations Hubble had to make in order to save the relativistic model suggested that it might be a "forced interpretation" of the data. This is a point that Hubble admitted himself, "by pressing the data to the limit of their tolerance, always in one direction, we might force the velocity-shifts into the framework of the surveys."[115]

Hubble also acknowledged that this procedure did not bring about the desired results. If it was assumed that the universe was expanding, then the data from the surveys should have indicated the nature and amount of curvature required to compensate for the discrepancies. However, this did not turn out to be the case, and Hubble was forced to argue that the problem must have resided in the data itself.

> Actually, no curvature can be found which exactly compensates for the apparent departures from uniformity in each of the surveys. Nevertheless, if we admit the presence of rather considerable systematic errors in the observations, it is possible to select a curvature which will more or less restore the homogeneity. Hidden errors of the necessary dimensions are by no means impossible in the very delicate investigations near the limits of a great telescope. They may be improbable, but they are not impossible. Therefore, the expanding universe can be saved by introducing a sufficient amount of spatial curvature.[116]

---

114 Hubble, 1937a, p. 59.
115 Hubble, 1936b, p. 201.
116 Hubble, 1937a, p. 60.

Furthermore, the amount of curvature Hubble found necessary to introduce in order to restore homogeneity presented even more problems. The amount was found to be positive and very large. The positive curvature indicated that the universe was closed and that it had a finite volume, although it had no boundaries in three-dimensional space. The size of the radius also indicated that the universe was very small, astronomically speaking. Hubble explained that this came as a complete surprise, because it implied that a large fraction, perhaps a quarter, of the universe was being explored with existing telescopes.[117] The restriction of time-scale and the small volume of the universe were viewed by Hubble as "strange and dubious" conclusions, but they were also a necessary consequence of the theory of a homogeneous expanding universe.[118]

At this point, Hubble knew two of the three necessary characteristics, derived from relativity theory, which he believed would fully describe the universe. The first was the nature of expansion or contraction. The universe was seen to be expanding; yet the rate of expansion had been slowing down for some time. The second characteristic was the nature of curvature, which was found to be large and positive. The third characteristic, the cosmological constant, was associated with the contents of the universe, and connected the contents with the two former characteristics. Investigating the nature of this latter quantity also resulted in what Hubble regarded as "dubious" conclusions.

The cosmological constant was associated with the contents of the universe. Although Hubble understood the characteristics of nebulae themselves, the nature of internebular space was unknown. Therefore he could not directly determine the numerical

---

117 Ibid.

value of the cosmological constant. However, with the aid of the known quantities of expansion and curvature, and the assumption that the mean density of the universe was greater than zero, Hubble could calculate upper and lower limits for the constant, and make a reliable estimate. Since the limits were narrow, Hubble believed he could confidently approximate the value of the constant to be about 4.5×10^{-18} (years)$^{-2}$.[119]

With these calculated characteristics, Hubble could then identify the type of universe that represented the actual universe. In addition to the fact that it did not accommodate red-shifts, the Einstein model could also be ruled out because the cosmological constant was calculated to be positive and larger than it would be in the Einstein universe. Hubble assigned the universe to a particular class of monotonic or ever-expanding universes. The combined characteristics allowed Hubble to further identify the universe, within the class, as a unique model known as the "ever-expanding universe of the first kind."[120] This was a model associated with Lemaître, and later, Eddington and others.

In this model, the universe would expand without reversal, although the rate of expansion would decrease over time. The radius would increase from zero to infinity. At each moment the model would be homogeneous, but the mean density would also diminish over time as the average distance between nebulae increased. Eventually, a state of complete isolation would be reached and the universe would approximate the de Sitter

---

118 Hubble, 1937b, p. 12.
119 Hubble, 1937a, p. 61.
120 Ibid., p. 62.

configuration - a universe so vast and thinly populated, that the nebulae are completely isolated in space.[121]

However, the density of this universe presented another problem for Hubble. The size restriction imposed by the curvature indicated a very massive universe. Hubble calculated the mean density to be on the order of a thousand times that which could be accounted for by observations. If the theory was valid, Hubble explained that the universe would have to be "packed with matter to the very threshold of perception." If such quantities of matter existed in internebular space in its familiar forms, it would absorb light. However, there was no trace of space absorption detected in the spaces between nebulae.

This forced Hubble to speculate about the presence of additional matter in an unlikely, invisible form, and to admit that the ever-expanding model seemed rather dubious. Observations did not rule out the model, but it suggested a forced interpretation of the data.[122] Although Hubble recognized discrepancies with the data, he once again emphasized the uncertainties, "if the estimations of density were completely reliable, a radius of curvature of the necessary dimensions would be ruled out by the evidence. But so definite a solution is probably unwarranted. The crucial data are surrounded by uncertainties."[123]

Thus, the investigation of the nature of red-shifts and the structure of the universe presented Hubble with a dilemma. The data could be fitted to support either of the two interpretations, which indicated two quite different types of universes. If red-shifts were

---

121 Hubble, 1937a, p. 62.
122 Ibid., 63.
123 Hubble, 1936b, p. 201.

not velocity-shifts, then the interpretation of the data was economical, simple, and plausible. There was no restriction of the time-scale, no trace of spatial curvature, no limitation of spatial dimensions, and no problem of internebular material. The observable region was thoroughly homogeneous, but it was an insignificant fraction of the whole that could not indicate the nature of the universe.[124] Hubble might have accepted this model if it were not for the fact that, in his opinion, no other satisfactory explanation of red-shifts was known, except for the familiar interpretation of red-shifts as velocity-shifts.[125]

If red-shifts were velocity-shifts, the universe would be closed, small, and dense. The universe would be rapidly expanding, but the rate of expansion would have decreased over a long period of time. A large fraction of the volume and temporal history of the universe would have been explored already with existing telescopes. According to Hubble, the disturbing features associated with this interpretation of the data were all introduced by the recession factor, that is, by the initial interpretation of red-shifts as velocity-shifts. The departures from the linear law of red-shifts, the departures from uniform distribution, the curvature necessary to restore homogeneity, the excess material demanded by the curvature, and the restricted time-scale and spatial dimensions were all just the recession factor in another form.[126] However, despite all the discrepancies that arose, Hubble continued to believe that the interpretation of recession was the only known satisfactory interpretation.

Therefore, the data presented as alternatives a curiously small-scale universe or a thereunto unrecognized principle of nature. Hubble summarized the dilemma as follows:

---

124 Hubble, 1937a, p. 63.
125 Hubble, 1937b, p. 12-13.
126 Hubble, 1937a, p. 63.

"The familiar interpretation leads to a strange and dubious universe; the alternative possibility, which some might call a strange and dubious interpretation, leads to a simple and familiar conception of a universe so vast that it can not be described in its totality."[127]

For Hubble, a definite choice would have to await the construction of larger telescopes with the ability to furnish data beyond the existing levels of uncertainty. This might have furnished the empirical evidence to settle the question once and for all, but I contend that it was not needed in order to see which model Hubble actually preferred. Hubble's evaluation of the cosmological models entailed by the two interpretations of red-shifts make it clear that Hubble in fact supported the assumption that red-shifts were velocity-shifts and that the universe was expanding.

First, let us consider the interpretation and corresponding model without recession. By Hubble's own admission, the assumptions and calculations in this case were more consistent with the observational evidence. Only the necessary corrections for the effects of red-shifts were applied to the data, and the result was a simple, plausible and familiar universe without any major discrepancies. When the recession factor was introduced, the consistency with the data dissolved and Hubble was force to follow the "uneconomical" procedure of introducing further assumptions to compensate for subsequent discrepancies that were all equivalent to the recession factor in one form or another. Even then, the result was a "strange and dubious" universe that still did not reach the same level of consistency with the data.

At this point, based on Hubble's own methodology, the expanding model given by relativistic theory should have been abandoned. Hubble claimed that he would

---

127 Hubble, 1937b, p. 13.

compare the results with respect to their simplicity and internal consistency, and the results of this comparison were clear, even to Hubble.[128] In his methodology, Hubble held observation in the highest regard, calling it an "impartial court of appeals." If a theory and its predictions did not pass the trials of this court, it was supposed to be revised or abandoned altogether. However, instead of rejecting the theory for its lack of consistency with observations, Hubble continually attempted to save the theory by introducing more assumptions and emphasizing the uncertainties in the data; thereby allowing him to fall back on the position that more observations would be needed to settle the dilemma. As Hetherington points out, "Hubble, the observational cosmologist, was reduced to arguing that the observations that conflicted with his vision of the universe might be incorrect."[129] Hubble's belief in an expanding universe and its corresponding interpretation of red-shifts was so strong, that he ignored, perhaps unknowingly, his own methodology.

Hubble's acceptance of an expanding universe, despite contrary evidence, was a product of his personal beliefs. In his philosophy of science, Hubble separated the realm of science and values, but in practice they tended to merge. It is also evident that theory had a special authority over observation in Hubble's work, even though he asserted that the two advanced together toward a common goal. Although the nonrecessional model was more consistent with the available data, the model presented major discrepancies for Hubble on a philosophical level, and these discrepancies were enough for Hubble to reject the idea.

The first and most obvious problem Hubble had with the nonrecessional model was the fact that it offered no permissible explanation of the red-shifts. As he saw it,

---

[128] Hubble, 1937b, p. 6.

"Many ways of producing such effects are known, but of them all, only one will produce large red-shifts without introducing other effects which should be conspicuous but actually are not found. This one known permissible explanation interprets red-shifts as due to actual motion away from the observer."[130] In order for the nonrecessional model to account for this phenomenon, a new, unknown principle of physics would have to be introduced. For Hubble this seemed like an unwelcomed ad hoc explanation.

Another major problem with this model was that in such a universe the observable region became insignificant. This fact challenged the very heart of Hubble's observational approach to cosmology, which required a significant homogeneous sample from which to extrapolate the nature of the universe as a whole. Without such a sample, Hubble could not make sound inferences about the nature of the universe. However, as Hetherington shows, not everyone accepted this approach. For instance, Harlow Shapely challenged this research program, arguing that there was a general unevenness of distribution, with the nebulae tending to cluster.[131] However, this violated one of Hubble's fundamental propositions, the cosmological principle, which assumed the universe to be homogeneous and isotropic. Although Hubble regarded this as a pure assumption, it was apparently a necessary characteristic for Hubble's universe.

The nonrecessional model also violated Hubble's second fundamental proposition, the correctness of general relativity, which he believed predicted a universe that was either expanding or contracting, and in which space was curved in the vicinity of matter. Therefore, the static model could not be a possibility because it would be

---

129 Hetherington, 1982, p.58.
130 Hubble, 1934a, p. 13.
131 Hetherington, 1982, p. 62.

unstable, and it exhibited no signs of motion or spatial curvature. Some astronomers, like Milne, questioned these propositions. Since the theory could not specify whether the instability would lead to expansion or contraction, Milne considered it to be incomplete.[132] He also considered the curvature of space to be an ad hoc hypothesis, introduced in order to account for discrepancies in the data.

For Hubble, however, the curvature of space was ensured by the theory. It was also necessary in order to satisfy other philosophical values, primarily homogeneity, that were essential to Hubble's scientific program. As for the completeness of the theory, Hubble recognized that the theory did not indicate either the direction of change or the rate of change, but he believed this could be discovered by observations. In fact, he argued that the empirical law of red-shifts could be accepted as evidence that the universe was expanding in a particular manner and at a known rate.[133]

As Hetherington points out, everyone disliked the idea of introducing ad hoc explanations into science, but they could not agree on what actually constituted an ad hoc explanation. "Hubble and most other cosmologist objected to the introduction of new principles of physics into their science. Zwicky did not. And Hubble had his own set of fundamental principles, some of which Milne viewed as ad hoc."[134] It was Hubble's fundamental principles that guided his observational approach, and therefore the methodology he actually used was strongly influenced by his personal values.

The accumulation of assumptions introduced to correct for discrepancies in the data and in order to save the theory was uneconomical. If Hubble strictly followed his

---

132 Ibid., p. 64.
133 Hubble, 1942, p. 124.
134 Hetherington, 1982, p. 67.

own expressed methodology and gave empirical evidence priority, then he should have abandoned the theory. The fact that he did not do this lends strong support to the idea that Hubble believed red-shifts indicated actual motion. Although he admitted the accumulation of assumptions seemed artificial, for Hubble it was a necessary circumstance to save his interpretation of the data. Instead of evaluating the theory against the available data, a process that led to discrepancies, Hubble assumed the theory to be correct, and when problems arose, he suggested the data were erroneous. By emphasizing uncertainties, Hubble could explain away some of the discrepancies, and make his additional assumptions more permissible. It is clear that philosophical values and theory played a major role in Hubble's scientific investigations and influenced his choice of universal models. Hubble may have considered observation to be an "impartial court of appeals," but it is evident that in some cases theory held jurisdiction over this court. Observation was key to Hubble's methodology, but in practice observation and analysis were guided by theory.

By rejecting the cosmological model that agreed with his data, Hubble ignored his own methodology. His reasons for doing this were mainly based on philosophical values, thereby contradicting his philosophy of science. For the purpose of this thesis, whether or not Hubble was aware of this contradiction is not as important as the fact that it demonstrates his persistent belief in a relativistic, expanding, homogeneous model of the universe. Since Hubble rejected the nonrecessional model, he also rejected the interpretation that red-shifts were not velocity-shifts. Likewise, by accepting the expanding model of the universe, it directly follows that he must have accepted the model's corresponding interpretation of red-shifts. Therefore, despite contrary evidence

and the skeptical standpoint exhibited in his writings, the conclusion can be made that Hubble interpreted red-shifts as the familiar Doppler-shifts, and accepted them as evidence of the expansion of the universe.

Conclusion

In this thesis, I have argued that in the early- to mid-twentieth century, various interpretations of the red-shifts that were present in the spectra of extra-galactic nebulae were offered. While the red-shifts apparently behaved, and were usually interpreted by many astronomers, as familiar Doppler shifts, these other interpretations included arguments that the light waves from distant nebulae were longer, that the vibrations were slower, or that the light quanta could lose energy as they traveled great distances though space.

Regardless of the interpretation however, by the 1930s, the investigation of red-shifts had become extremely important to cosmology. Depending on which interpretation was assumed, one would obtain a vastly different picture of the universe that we inhabit. Therefore, the interpretation of red-shifts was of fundamental importance to astronomers deciding on a cosmological model that explained the reality of the universe.

As one of history's greatest astronomers, and the leading authority in regards to red-shifts and extra-galactic nebulae in the period between the early 1920s and early 1950s, Edwin Hubble's beliefs on this important cosmological matter are crucial indicators for understanding how top scientists of the early-twentieth century viewed our universe. The rationale for his beliefs, and the ways in which he evaluated data and presented arguments, are good indications of how scientists viewed the relationship between theory and observations, and provide insight into how new scientific knowledge gained acceptance in a time when cosmology was becoming a respected scientific discipline in its own right.

Despite what he wrote to be insufficient empirical evidence for the interpretation, Edwin Hubble privately believed that the red-shifts in the spectra of extra-galactic nebulae exhibited the expansion of the universe. By adopting such a belief, he rejected his own scientific "court of appeals", observation, and thereby violated his own methodology of science. In doing so, he also refuted his own expressed philosophical views , at the core of which was the idea of science as progressive, cumulative, and objective.

Hubble attributed, as do several historians[135], much of his views about science to the influence of the Harvard historian of science, George Sarton. For Sarton, progress was of fundamental importance to history. In *The History of Science and the New Humanism*, written in 1931, he stated that, "human history is not truly significant unless we can describe the progress of mankind along a certain direction,"[136] and "in almost every case wherever there is progress, or the possibility of progress, that is due to science and its applications. I would never claim that science is more important than art, morality, or religion, but it is more fundamental, for progress in any direction is always subordinated to some form or other of scientific progress."[137]

In light of this, it is easy to understand why Sarton attributed such an indispensable importance to science, which he considered to be the only activity that is obviously and undoubtedly cumulative and progressive. Therefore, science, and scientific knowledge could be used as measures of the progress of civilization. For Sarton, our

---

135 See the works of Norriss Hetherington for example.
136 Sarton, 1931, p.10-11.
137 Ibid., p.20.

knowledge of science, or systematized positive knowledge, is the essential feature that distinguishes modern civilization from those that came earlier.

> Throughout the course of history, in every period, and in almost every country, we find a small number of saints, of great artists, of men of science. The saints of today are not necessarily more saintly than those of a thousand years ago; our artist are not necessarily greater than those of early Greece; they are more likely to be inferior; and of course, our men of science are not necessarily more intelligent than those of old, yet one thing is certain, their knowledge is at once more extensive and more accurate. The acquisition and systemization of positive knowledge is the only human activity which is truly cumulative and progressive. Our civilization is essentially different from earlier ones, because our knowledge of the world and of ourselves is deeper, more precise, and more certain, because we have gradually learned to disentangle the forces of nature, and because we have contrived, by strict obedience to their laws, to capture them and to divert them to the gratification of our own needs.[138]

Hubble regarded the cumulative and progressive attributes of science as remarkable, although they are bought at a price, namely, the strict limitation of the subject matter.[139] Here Hubble shows the additional influence of Norman Campbell regarding his views towards science. Campbell defined science as "the study of those judgments concerning which universal agreement can be obtained… According to the definition proposed, the things which science studies are very closely allied to those which make up the external world of nature."[140] Hubble combined these characteristics of science into his own philosophy, which appears in many of his works. In his manuscript

---

138 Sarton, 1927, p. 3-4.
139 Hubble, 1954, p. 40.
140 Campbell, 1952, p. 27.

on *English Science in the Renaissance*, written in 1937, Hubble provides a clear summary of these views:

> Science deals with judgments on which it is possible to obtain universal agreement. These judgments do not concern individual facts and events, but the invariable association of facts and events known as the laws of science. Agreement is secured by observation and experiment – impartial courts of appeal to which all men must submit if they wish to survive. The laws are grouped and explained by theories of ever increasing generality. The theories at first are ex post facto- merely plausible interpretations of existing bodies of data. However, they frequently lead to predictions that can be tested by experiments and observations in new fields, and, if the interpretations are verified, the theories are accepted as working hypotheses until they prove untenable. The essential requirements are agreement on the subject matter and the verification of predictions. These features insure a body of positive knowledge that can be transmitted from person to person, and that accumulates from generation to generation.
>
> It should be emphasized that the necessity for universal agreement completely bars science from the great world of values, the world in which we spend most of our lives. Judgments of beauty and reverence, for instance, are purely personal judgments. These are not external tests of validity; there is no calculus of values. Wisdom cannot be directly transmitted, and does not readily accumulate through the ages.[141]

For Sarton, as well as Hubble, the main purpose of science is the discovery of truth. The pursuit of science is characterized by disinterested curiosity, and the success in ascertaining objective truths is measured by the degree to which personal desires are suppressed and by the amount of significant data that can be isolated.[142] Science is an

---

141 Hubble, 1937d. Quoted in Hetherington, 1988, p. 5.
142 Hubble, 1954, p.41-42.

attempt to explain nature in its own terms, and to evidence its unity, wholeness, and congruency. Sarton maintained that, "the very fact that science exists, and that it does not simply exist, but 'works' and pays fantastic dividends, proves the consistency of nature in a thousand ways."[143]

The unity of nature was a main postulate of science for Sarton, and was indirectly confirmed for him by the development of knowledge. "It is clear enough that if there were no unity in nature, if the universe were not a cosmos, but a chaos, if there were no regularities, no laws, but only an erratic succession of miracles, there would be no occasion for scientific research, no possibility of scientific progress."[144] This postulate, the uniformity of nature, was also adopted and strongly advocated by Hubble. In fact, it is closely related to what Hubble called the cosmological principle, which was one of the fundamental principles from which he began his investigations.

The cosmological principle supposes that the universe is homogeneous and isotropic - very much the same everywhere and in all directions. This ensures that there is no favored position in the universe, and that the universe appears the same to all observers, regardless of their location in it. The influence of this principle on Hubble's work can be seen in his observational approach to cosmology, which consists of inferring the nature of the universe as a whole from the observed characteristics in the sample of the universe available for inspection. In order to make sound inferences, the observable portion of the universe has to be considered a sufficient representation of the whole. This is a belief Hubble would maintain despite any amount of evidence to the contrary,

---

143 Sarton, 1931, p. 24.
144 Sarton, 1927, p. 30-31.

because without it, as Sarton put it, "there would be no occasion for scientific research and no possibility of scientific progress."[145]

In the methodology Hubble outlines, he values observation and experimentation above all else, even considering this empirical approach an "impersonal court of appeal." Theories are then derived to explain observations and to make further predictions that could be tested with more observations. A theory is then evaluated based on its consistency and the verification of its predictions. No theory is to be regarded as sacred, and if it fails to meet the tests of observation, it has to be modified or abandoned.

Sarton also argues that the central method of science is the experimental method, which he says "consists essentially in arranging things in such a way that the answers to our questions are given by nature itself, thus apparently eliminating ourselves."[146] Similar to Hubble's account, Sarton explains the empirical process as follows:

> The study of a group of phenomena leads to the formulation of certain laws. When one tries to apply these laws to the explanation of other phenomena either one succeeds and their validity is confirmed, or one fails and other laws are eventually discovered which account for all the phenomena. This has worked so well that when we have failed to give a consistent explanation of all facts involved, we have hardly been dismayed by our failure; we have simply concluded that our knowledge was insufficient, and that as soon as it was sufficiently developed we would find a means of conciliating all the data. In other words the success of this method has been so marvelous that we are more ready to attribute our relatively few failures to our ignorance or to the radical infirmity of our minds than to an inconsistency inherent in nature.[147]

---

145 Ibid.
146 Sarton, 1931, p. 25-26.

Thus, it is clear, from his written works, that Hubble held views pertaining to science similar to Sarton, and that he portrayed science as a cumulative and progressive human activity that followed an empiricist methodology to the end of ascertaining objective truths about nature.[148] However, this realization alone is an insufficient assessment of Hubble's work and career. One must ask why, if Hubble was ostensibly committed to an empiricist framework, did he consistently lean towards positing an expansionist understanding of the universe when the observations, his "impartial court of appeals," were more consistent with a static model. The answer is that, whether Hubble was aware of it or not, he did not practice what he preached. In practice, Hubble did not adhere to his own empirical methodology, and instead allowed his own theoretical presuppositions to guide his observations and the interpretations of his data.

In order to explain this apparent paradox, it is therefore necessary to look beyond both Hubble's own theoretical and methodological discussions and these of other writers, like Sarton, whom Hubble explicitly followed. Instead, one must examine Hubble's writings anew in order to isolate the preconceptions that guided his empirical work, and

---

147 Ibid., p. 24.

148 The idea of cumulative and progressive science had been long and widely held; see Zilsel (2000) who traces the idea back to Francis Bacon, after whom it became a common view widely held by scientists. (A shortened version of Zilsel's article published in 2000 was published under the title "The Genesis of the Concept of Scientific Progress" in the *Journal of the History of Ideas*, 1945, 6, p. 325-349.) However, probably noone in the early 20th-century discussed and embodied these characteristics of science in their work more than George Sarton. Although his views may have been common, his works were obviously influential to many, as is indicated by Hubble's citations of them. Hubble was not alone at this by a long shot: "An indication of the impact of Sarton's works is given by data accumulated from 1955 to 1984 in the Institute for Scientific Information's (ISI's) Science Citation Index (SCI), Social Sciences Citation Index (SSCI), and from the Arts and Humanities Citation Index (A&HCI). Most of Sarton's works were published well before the earliest date for which SCI data are available. Thus, these works may have passed through their peak citation years before they were indexed in SCI. Even so, his three-volume *Introduction to the History of Science*, for example, has been cited over one hundred and fifty times." Garfield, 1985, p. 107. This, among his other accomplishments in the field, may be why Sarton gained the reputation as the father of the history of science.

to demonstrate how these preconceptions in fact influenced his interpretations of scientific data.

When Hubble concluded his evaluation of the data, he maintained that it was possible to explain his results on the basis of either a static homogeneous model of the universe with some unknown cause for red-shifts, or an expanding homogeneous model with the introduction of effects from spatial curvature that may seem unexpectedly large but not impossible.[149] The degree of uncertainty expressed in regards to the latter model was necessary, because, as we have seen, although the expanding model was not ruled out by the data, the static model was definitely more consistent with the results obtained from the investigations.

Therefore, Hubble believed that the data presented as alternatives a curiously small-scale universe or an unrecognized principle of nature. He summarized his dilemma: "The familiar interpretation leads to a strange and dubious universe; the alternative possibility, which some might call a strange and dubious interpretation, leads to a simple and familiar conception of a universe so vast that it can not be described in its totality."[150]

According to Hubble, a definite choice between these two options would have to await the construction of larger telescopes with the ability to furnish data beyond the existing levels of uncertainty. However, I contend that despite his claim, we can understand which model Hubble actually preferred. As this thesis has demonstrated, Hubble's evaluation of the cosmological models entailed by the opposing interpretations of red-shifts makes it clear that Hubble in fact supported the assumption that red-shifts were velocity-shifts and that the universe was expanding.

---

149 Hubble and Tolman, 1935, p. 335.

Although the static cosmological model was more consistent with the empirical data collected by Hubble, which should have made it more appealing to him, the model proved to be incompatible with some of his fundamental beliefs. This model prescribed a universe that was so vast, that it could not be described in its totality. Therefore, our observable portion would be an insufficient sample and could not provide any positive knowledge about the universe as a whole. We have seen that regarding our observable region of space as a sufficient sample of the universe in order to make inferences about the whole was a necessity for Hubble's scientific approach. Consequently, this characteristic of the universe entailed by a static model directly opposed one of Hubble's fundamental beliefs.

The necessity of a sufficient sample was a product of Hubble's assumption of the cosmological principle, that is, the uniformity of nature without which, he believed, science could not be carried out. This theoretical presupposition is significant, because it influenced Hubble's work at every stage. The necessity for a sufficient sample was a proposition from which Hubble began his research, but one can also see the influence of the cosmological principle on Hubble in terms of obtaining and evaluating observations.[151]

By utilizing the cosmological principle, Hubble assumed upper limits of absolute luminosity for stars and nebulae. This allowed him to observe distant sources and estimate their distances by comparing their apparent luminosity with the luminosities of

---

[150] Hubble, 1937b, p.13.

[151] The principle of the uniformity of nature was a fundamental presupposition that Hubble supported throughout his career. For instance, in "NGC 6822, a Remote Stellar System," written in 1925, Hubble contends, "The principle of the uniformity of nature thus seems to rule undisturbed in this remote region of space. This principle is the fundamental assumption in all extrapolations beyond the limits of known and

sources whose distances were already determined. Of course, there is no guarantee that there is an upper limit to absolute luminosity, but, assuming the uniformity of nature, a star of a certain type and with a certain brightness should have the same characteristics everywhere in the universe. Therefore, assuming similarities between stars and nebulae of certain types allowed Hubble to make even greater estimates of distance than were possible with previous methods.

The influence of the cosmological principle can also be seen on Hubble's evaluation of data. For example, we have seen that a red-shift, regardless of its interpretation, decreases the energy of light quanta by increasing the wave-length. Therefore, in either the case of actual motion or not, the data must be corrected for these energy effects. However, interpreting red-shifts as velocity-shifts introduces the need for further corrections. Since actual motion away from an observer lengthens the path that light quanta travel from a nebulae, the rate at which the quanta reach the observer is reduced by a certain factor. This factor is known as the number effect, and operates only in the case of actual motion.

When Hubble applied the necessary corrections for energy effects alone, the results were consistent with his observations. The law of red-shifts remained linear and the distribution of nebulae was uniform and homogeneous. It was possible to approximate the universe with the Einstein model or a homogeneous expanding model in which the curvature of space and the rate of expansion were inappreciable. This interpretation of the data, according to Hubble, was direct, economical and very simple;

---

observable data, and speculations which follow its guide are legitimate until they become self-contradictory." Hubble, 1925, p. 432.

however, it meant that the observable region was an insignificant fraction of the universe.[152]

When Hubble analyzed the data under the interpretation of velocity-shifts and made the corrections for the recession factor, the results appeared much more inconsistent, especially with respect to his fundamental beliefs. For instance, Hubble found that the distribution of nebulae was not uniform, and the density of nebulae increased outward in all directions. The observable region appeared to be non-homogeneous, indicating a favored position for the observer. This conflicted with Hubble's belief in the cosmological principle, which assumed the homogeneity of the universe and prevented any observer from occupying a favorable position within it.

Hubble's confidence in the truth of the cosmological principle allowed him to assume that the departures from uniformity and homogeneity were not real, and thus to search for neglected factors to compensate for the apparent discrepancies. In order to compensate for the discrepancies, Hubble introduced spatial curvature, which was prescribed by another theoretical presupposition that he held as a fundamental belief, namely, the correctness of general relativity. Hubble claimed that general relativity predicts a universe that is unstable, either expanding or contracting, and in which space is curved in the vicinity of matter. Guided by these beliefs, Hubble proposed to posit expansion, and then to postulate spatial curvature in order to compensate for the departures from homogeneity. Although Hubble acknowledged that this procedure

---

152 Hubble, 1936a, p. 553.

seemed artificial, it offered, in his opinion, the only means of preserving the theory of a homogeneous, expanding universe:[153]

> To the observer the procedure seems artificial. He has counted the nebulae to various limits, applied only the corrections that are necessarily required (energy-corrections), and derived the quite plausible result of uniform distribution. Now, in testing the relativistic theory, he introduces a new postulate, namely, spatial curvature, in order to compensate the discrepancies introduced by the first. The accumulation of assumptions is uneconomical, and the justification must be sought in the general background of knowledge. The outstanding argument is the fact that velocity-shifts remain the only permissible interpretation of red-shifts that is known at the present time.[154]

Although Hubble defended the need for this artificial procedure, he was forced to acknowledge that it did not bring about the results he desired. If the universe was assumed to be expanding, then the data should have indicated the nature and amount of curvature required to compensate for the discrepancies. However, this did not turn out to be the case, and Hubble then began to argue that the problem must be a feature of the data itself:

> Actually, no curvature can be found which exactly compensates for the apparent departures from uniformity in each of the surveys. Nevertheless, if we admit the presence of rather considerable systematic errors in the observations, it is possible to select a curvature which will more or less restore the homogeneity. Hidden errors of the necessary dimensions are by no means impossible in the very delicate investigations near the limits of a great telescope. They may be improbable, but

---

153 Hubble, 1937b, p. 11.
154 Hubble, 1937a, p. 59.

they are not impossible. Therefore, the expanding universe can be saved by introducing a sufficient amount of spatial curvature.[155]

Hubble's belief that velocity-shifts were the only permissible interpretation of red-shifts can also be attributed to his philosophy of science. Although the effects of red-shifts could be explained in other ways, such as a possible gravitational drag on light, the cause of red-shifts would still be unknown. Since the fundamental cause of red-shifts remained unknown, the result would be the introduction into science of an unrecognizable principle of nature. By the early twentieth century, astronomers were familiar with velocity or Doppler shifts in relation to stars, but, as we have seen from Hubble's work, applying this interpretation to the red-shifts of nebulae was not as straight forward. Similar to Sarton, Hubble was more inclined to attribute his insufficient amount of empirical proof for the interpretation to "our ignorance or to the radical infirmity of our minds than to an inconsistency inherent in nature."[156] In Hubble's opinion, the inconsistencies were more likely to be found in the data rather than nature. Thus, by emphasizing uncertainties or possible errors in the data, he could privately maintain his belief in the expanding universe, while publicly requesting the need for more accurate observations in order to settle the issue once and for all.

Overall, it is clear that Hubble's methodological prescriptions do not translate cleanly into the relationship between theory and practice seen in his work. The fact that such an eminent astronomer as Hubble could not live up to the inductive, observational framework espoused by him and so many others calls into question the viability of this

---

155 Hubble, 1937a, p. 60.

methodology. More specifically, the ideal of pure, impartial objectivity in scientific practice appears to be unreasonable, and the attempt to locate science beyond the realm of subjective, social human activity cannot be entirely accurate.

We have seen that in the philosophy espoused by those like Hubble and Sarton, science is viewed as the pursuit of objective truths about nature. Knowledge is obtained by careful experimentation and observation. By these empirical means, nature's truths reveal themselves to the observer in the form of natural, indisputable "facts." Theories are then devised that attempt to accumulate and explain these phenomena more generally. In order to validate or refute a theory, observation and experimentation are again employed to test predictions offered by a theory. If a theory survives the further tests, it is then regarded as a working hypothesis and is added to the accumulative body of positive scientific knowledge.

However, numerous historians, philosophers, and sociologists of science have recently argued that scientific activity must be seen as a fundamentally social activity. In addition, they argue that knowledge or truth is not some objective entity that can be ascertained by careful examination of nature, but, on the contrary, they assert that knowledge is something that is socially constructed. Knowledge or truth is not absolute; rather, it is what a community collectively considers to be true. In other words, it is a social construct. In *A Social History of Truth*, Steven Shapin states:

> …what we know about the world is arrived at, sustained, and recognized through collective action…no single individual can constitute knowledge: all the individual can do is offer claims,

---

156 Sarton, 1931, p. 24.

> with evidence, arguments, and inducements, to the community for its assessment. Knowledge is the result of the community's evaluations and actions, and it is entrenched through the integration of claims about the world into the community's institutionalized behavior.[157]

Interpreting knowledge in this way prohibits any sense of objectivity from being applied to it. Shapin contends: "Communities making truth-judgments mean to distinguish statements or beliefs which correspond to reality from those which do not, and as they do so they create an automatic bias in favor of their own stock of current knowledge."[158] Therefore, knowledge, or the pursuit of knowledge, cannot possibly be considered an objective activity.

David Bloor offers an argument similar to Shapin in his book *Knowledge and Social Imagery*. In opposition to the empirical philosophy propounded by Hubble, Bloor provides an alternative that he calls the strong programme in the sociology of knowledge, and that is defined by the four principles of causality, impartiality, symmetry, and reflexivity[159]:

> 1.) It (the strong programme) would be causal, that is, concerned with the conditions which bring about belief or states of knowledge. Naturally there will be other types of causes apart from social ones which will cooperate in bringing about belief.
>
> 2.) It would be impartial with respect to truth and falsity, rationality or irrationality, success or failure. Both sides of these dichotomies will require explanation.

---

157 Shapin, 1994, p. 6.

158 Shapin, 1994, p. 4.

159 Bloor, 1991, p. 7. Although many recent authors have dealt with the topic of how scientific knowledge is obtained, Bloor is selected here because his position represents the opposite side of the spectrum, so to speak, in relation to Hubble's views. Whereas Hubble believed knowledge or truth was an objective entity that could be obtained by careful examination of nature, Bloor believes knowledge is socially constructed dependent solely upon subjective factors.

3.) It would be symmetrical in its style of explanation. The same types of cause would explain, say, true and false beliefs.

4.) It would be reflexive. In principle its patterns of explanation would have to be applicable to sociology itself. Like the requirement of symmetry this is a response to the need to seek for general explanations. It is an obvious requirement of principle because otherwise sociology would be a standing refutation of its own theories.

Bloor argues that an inductivist methodology, like the one espoused by Hubble, which stresses the emergence of theories out of an accumulation of observations can never completely explain all of the diversity of actual scientific practice.[160] The theory that underlies these ideas is a goal-directed or teleological vision of the advance of knowledge. He also contends that if it is assumed that truth, rationality and validity are our natural goals and the direction of certain natural tendencies that we are endowed with, and if it is assumed that we are rational animals so that we naturally reason justly and hold to the truth whenever possible, then beliefs that are true require no special comment or explanation. For these beliefs, their truth is all the explanation that is needed as to why they are believed.[161] On the other hand, Bloor contends, "this self-propelling progress towards truth may be impeded or deflected and here natural causes must be located. These will account for ignorance, error, confused reasoning, and any impediment to scientific progress."[162]

Hubble and Sarton held such a teleological view of knowledge - the cumulative and progressive pursuit of truth - and Bloor's description can be seen in their exposition

---

160 Bloor, 1991, p. 9.
161 Bloor, 1991, p. 11.
162 Bloor, 1991, p. 11.

of their method.[163] When the empirical method concludes an experiment with results considered to be valid, the reason why that method worked is not questioned. On the other hand, when the method fails, it is assumed that there must be a problem with the scientist or instrumentation and a cause for the error is then looked for.

Bloor then asks, how does this teleological model of knowledge relate to the tenets of the strong programme? His answer is that it clearly violates them in a number of ways. First, the principle of causality is violated. Since causes can only be located for error, the sociology of knowledge found in the teleological model is confined to the sociology of error. Furthermore, the principles of symmetry and impartiality are also violated. "A prior evaluation of the truth or rationality of a belief is called for before it can be decided whether it is to be counted as self-explanatory or whether a causal theory is needed. There is no doubt that if the teleological model is true then the strong programme is false."[164]

However, Bloor maintains that the teleological view does not have to be demonstrated to be false in order for the strong programme to proceed. Instead, he argues that:

> It is more sensible to look at matters from the other way around. It is unlikely that any decisive, independent grounds could be adduced 'a priori' to prove the truth or falsity of such major metaphysical alternatives. Where objections and arguments are proposed against one of the theories it will be found that they depend on and presuppose the other, and so beg the question at issue. All that can be done is to check the internal consistency of the different theories and then see what happens when practical research and theorizing is based upon them. If their truth can be

---

163 See footnote 143.

decided at all it will only be after they have been adopted and used, not before. So the sociology of knowledge is not bound to eliminate the rival standpoint. It only has to separate itself from it, reject it, and make sure that its own house is in logical order.[165]

Despite this claim, Bloor does offer a reason why the strong programme should be accepted over its rival:

> What may be said, however, is that the strong programme posses a certain kind of moral neutrality, namely the same kind as we have learned to associate with all the other sciences. It also imposes on itself the need for the same kind of generality as other sciences. It would be a betrayal of these values, of the approach of empirical science, to choose to adopt the teleological view.[166]

Another aspect of the philosophy of science that Hubble followed is its individualistic character with respect to obtaining knowledge. In other words, by applying an empirical methodology to the study of nature an individual can obtain knowledge in isolation. Bloor attributes this characteristic to a theory he refers to as empiricism and challenges the claim that knowledge is acquired in this way.[167] In doing so, Bloor does acknowledge that an individual's experience may contribute to the

---

164 Bloor, 1991, p. 12.

165 Bloor, 1991, p. 12.

166 Bloor, 1991, p. 13.

167 Bloor describes the empirical theory as functioning in the following way: "social influences produce distortions in our beliefs whilst the uninhibited use of our faculties of perception and our sensory-motor apparatus produce true beliefs. This praise for experience as a source of knowledge can be seen as encouraging individuals to rely on their own physical and psychological resources for getting to know the world. It is a statement of faith in the power of our animal capacities for knowledge. Give these full play and their natural, but causal, operation will yield knowledge tested and tried in practical interaction with the world. Depart from this path, rely on one's fellows, and one will be prey to superstitious stories, myth and speculation. At best these stories will be second-hand belief rather than first-hand knowledge. At worst the motives behind them will be corrupt, the product of liars and tyrants." Bloor acknowledges that this is a version of Bacon's warning to avoid the Idols of the Market Place and the Theatre. Bloor, 1991, p. 14.

formulation of knowledge, but he questions the extent of this contribution and emphasizes the social influences inevitably acting upon it.

> These aspects of knowledge which each of us can and has to furnish for himself may be adequately explained by this type of model. But how much of man's knowledge, and how much of his science is built up by the individual relying simply on the interaction of the world with his animal capacities? Probably very little…Does not individual experience, as a matter of fact, take place within a framework of assumptions, standards, purposes and meanings which are shared? Society furnishes the mind of the individual with these things and also provides the conditions whereby they can be sustained and reinforced… As well as the individual's sensory experience of the natural world, there is, then, something that points beyond that experience, that provides a framework for it and gives it a wider significance. It fills out the individual's sense of what that overall Reality is, that his experience is experience of.[168]

Bloor believes that the knowledge of a society is not so much the collective experience of its individual members, or the sum of what he calls their "animal knowledge," but rather their collective vision or visions of reality:[169]

> Thus the knowledge of our culture, as it is represented in our science, is not knowledge of a reality that any individual can experience or learn about for himself. It is what our best attested theories, and our most informed thoughts tell us is the case, despite what the appearances may say. It is a story woven out of the hints and glimpses that we believe our experiments offer us. Knowledge then, is better equated with Culture than Experience…(therefore) the empiricist theory is implausible as a description of what we in fact count as our knowledge.[170]

---

168 Bloor, 1991, p. 15.
169 Bloor, 1991, p. 15-16.
170 Bloor, 1991, p. 15-16.

Applying this theory to the case of Hubble allows us to gain some understanding as to why he privately continued to support the idea of an expanding universe, despite the contrary picture offered by his data. His decision was influenced by his fundamental beliefs that were regarded as true by the scientific community to which he belonged. Despite his efforts, in theory if not in practice, to distinguish personal values from empirical research, the former still managed to affect his work and influence his beliefs. Once again, the validity of the idea of objective science can be called into question.

The historian of science, Norriss Hetherington, argues that the ideal, or idol, of objective science was an idea whose time had come in the aftermath of the First World War.[171] Hetherington maintains that the Great War caused so much destruction to Western civilization's material and intellectual culture that salvation was sought in science, despite the role of science in the destruction of previous values. Objective science therefore became "a necessary totem for a civilization that had lost faith in its other conventions," and "the image of objective science shone as a beacon lighting man's way to true progress."[172]

Hetherington argues that the ideal of objectivity came to characterize science and to distinguish it from, and raise it above, all other human activity; however, science remains a human activity, subject to inevitable human bias.[173] This bias can range from the occasional deliberate fraud to the more common influence of preconception on

---

171 Hetherington, 1988, p. 3.

172 Hetherington, 1988, p. 4-6. Hetherington regards this search for salvation in science as a paradox, since changes in science had played a part second only to that of the Great War in the destruction of European civilization. For instance, He argues that the "Darwinian evolution had begun the weakening of civilization's intellectual underpinnings before the war, and Einsteinian relativity contributed to the feeling of lost directions and values after the war." p. 3.

perception. Furthermore, not only can this bias affect the interpretation of data, but the collection of data as well:

> Preconception also helps determine the vigor and thoroughness with which data is gathered. It is difficult to determine when an experiment or observation has been completed, when all the errors have been weeded out and only the truth remains. One indication of the status of an experiment or observation is its degree of agreement with the expected result. In the face of disagreement, a scientist's natural response is to check for errors. Should there be early agreement between observation and preconception, a scientist may be lulled into premature complacency.[174]

A parallel to this can be drawn in the case of Hubble, where he emphasizes possible errors and asserts the need for further observations.

Therefore, however attractive and assuring the image of objective science is, Hetherington believes that the picture of objective science and civilization is flawed and does not offer a perfectly accurate representation of reality.[175] "The image requires modification, qualification, and perhaps even replacement, no matter how psychologically comforting and culturally necessary has the idol been for its guarantee of human progress."[176] I would have to agree with Hetherington on this point, and this thesis can be viewed as an example of a failure of objectivity in science. It shows that even great scientists, such as Edwin Hubble, are not immune to preconceptions and the influence of fundamental beliefs despite their efforts, both in theory and practice, to

---

173 Hetherington, 1988, p. 111.
174 Hetherington, 1988, p. 112. See also Galison, 1987.
175 Hetherington, 1988, p. 6.
176 Hetherington, 1988, p. 111.

separate the realm of science from the realm of values in an attempt to portray an image of science as the only truly progressive, cumulative, and objective human activity.

Although one is led to conclude, from the case of Hubble, that scientific activity must be seen as inevitably partly a social endeavor, this does not mean that science should be degraded, or relegated to the realm of the purely subjective in a demeaning manner. Pure objectivity may not be obtainable in science, and it might be naïve to contend otherwise in light of the many instances of failed objectivity; however, this realization should not be used to attack the importance or credibility of science as a whole. Science may not be completely objective, but efforts can and are taken to make it as objective as possible. Likewise, the concept of objectivity should not be interpreted as rigidly as Hubble and other empiricist scientist interpret it.

On the one hand, science is similar to other forms of human activities that involve the interplay of humans and nature. Hubble does not address this aspect of science when he quotes from Sarton, despite the fact that Sarton does in fact discuss it elsewhere in his work. In addition to emphasizing the empirical aspects of scientific activity, Sarton argues that science as a whole must be seen as a humanistic endeavor:

> Nothing could be more foolish than to oppose the study of nature to the study of man, because in both cases we have to deal with the same fundamental dualism--- man vs. nature… On the other hand the study of nature is of necessity a human study of it. However objective it may be, and scientists try to make it as objective as possible, it remains anthropocentric. Personal idiosyncrasies can and must be eliminated, but humanity itself cannot. Science is nothing but the human mirror of nature.[177]

---

[177] Sarton, 1931, p. 24-25.

On the other hand, the uniqueness of science should also be recognized. Science is cumulative in the sense of being a body of knowledge held in common that scientists build upon in successive generations. For example, it is possible for a modern artist to create an artistic masterpiece without really knowing the works of prior artists. Yet scientific knowledge is built upon the understanding of nature held by those who came before. Although modern art can be based on and regarded as equal to previous art in terms of beauty, it does not contribute to one universal understanding of what beauty is. In contrast, modern science stems from previous scientific knowledge and attempts to increase our understanding of nature and Mankind's place within it. In addition, even though our scientific knowledge may never fully attain its goal of a complete and accurate understanding of nature, the pursuit of this endeavor is admirable in its own right. As Sarton explains,

> It is possible, nay probable, that we shall never reach the truth itself, but the steady application of the scientific spirit in all its forms will bring us nearer and nearer to it. The most valuable portion of our experience is not our scientific knowledge but our constant endeavor to obtain it. Next to nature there is nothing more wonderful than man's gradual understanding of it. Next to truth, there is nothing more impressive than man's patient effort to reach it irrespective of consequences, simply because his own being exacts it from him. That is undoubtedly a part of man, and perhaps the best part of him; the noblest aspect of his humanity.[178]

Edwin Hubble's work exemplifies this view of science as a human activity but with its own unique characteristics. Although Hubble developed his contributions to

science in a manner not keeping with his stated methodology and philosophy of science, his contributions were nonetheless very important. Not only did he, figuratively, expand the boundaries of our universe, but he also helped to expand our knowledge of it. Therefore, because of his momentous contributions to science, it is justifiable to ascribe to Hubble the well-deserved title of one of history's greatest astronomers.

---

178 Sarton, 1931, p. 29.

Bibliography

Primary Sources

-Hubble, 1925. = Hubble, Edwin. "N.G.C. 6822, a Remote Stellar System." *Astrophysical Journal*, 62: 409-433.

-Hubble, 1929. = Hubble, Edwin. "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae." *Proceedings of the National Academy of* Sciences, 15, no. 3 (March 1929), 168-173.

-Hubble, 1934a. = Hubble, Edwin. *Red-Shifts in the Spectra of Nebulae being The Halley Lecture delivered on 8 May 1934*. Oxford: The Clarendon Press, 1934.

-Hubble, 1934b. = Hubble, Edwin. "The Distribution of Extra-Galactic Nebulae." *Astrophysical Journal*, 79 (1934), 8-76.

-Hubble, 1936a. = Hubble, Edwin. "Effects of Red Shifts on the Distribution of Nebulae." *Astrophysical Journal*, 84 (1936), 517-554.

-Hubble, 1936b. = Hubble, Edwin. *The Realm of the Nebulae.* Oxford: Oxford University Press, 1936.

-Hubble, 1937a. = Hubble, Edwin. *The Observational Approach to Cosmology.* Oxford: The Clarendon Press, 1937.

-Hubble, 1937b. = Hubble, Edwin. *Our Sample of the Universe.* Washington D.C., Carnegie Institution of Washington, 1937.

-Hubble, 1937c. = Hubble, Edwin. "Red Shifts and The Distributions of Nebulae." *Monthly Notices of the Royal Astronomical Society*, v. 97, (1937), 506-513.

-Hubble, 1937d. = Hubble, Edwin. Manuscript on *English Science in the Renaissance*, 1937, Hubble Collection, box 2, Huntington Library, San Marino California.

-Hubble, 1942. = Hubble, Edwin. "The Problem of the Expanding Universe." *Annual Report of the Board of Regents of The Smithsonian Institution 1942*, 119-132.

-Hubble, 1953. = Hubble, Edwin. "The Law of Red-Shifts." *Monthly Notices of the Royal Astronomical Society*, 113, no. 6 (1953), 658-66.

-Hubble, 1954. = Hubble, Edwin. *The Nature of Science, and Other Lectures.* San Marino, California: Huntington Library, 1954.

-Hubble and Humason, 1931. = Hubble, Edwin and Milton L. Humason. "The Velocity-Distance Relation Among Extra-Galactic Nebulae." *Astrophysical Journal*, 74 (1931), 43-80.

-Hubble and Tolman, 1935. = Hubble, Edwin and Richard C. Tolman. "Two Methods of Investigating the Nature of the Nebular Red-Shift." *Astrophysical Journal*, 82 (1935), 302-337.

-Mayall to Hubble, 1934. = Mayall, Nicholas U. to Edwin Hubble, 6 March 1934. Hubble Collection, box 18, Huntington Library, San Marino, California.

-Mayall to Hubble, 1937. = Mayall, Nicholas U. to Edwin Hubble, 16 March 1937. Hubble Collection, box 18, Huntington Library, San Marino, California.

-Slipher, 1914. = Slipher, V. M. "Spectrographic Observations of Nebulae." *Publications of the American Astronomical Society*, 3 (1918), 98-100.

-Tolman, 1934. = Tolman, R. C. *Relativity, Thermodynamics, and Cosmology*. Oxford: Clarendon Press, 1934.

-Wirtz, 1924. = Wirtz, C. W. "De Sitters Kosmologie und die Radialbewegungen der Spiralnebel." *Astronomische Nachrichten*, 222, 21-26.

Secondary Sources

-Bloor, 1991. = Bloor, David. *Knowledge and Social Imagery*. 2d ed. Chicago and London: University of Chicago Press, 1991.

-Brush, 2001. = Brush, Stephen G. "Is the Earth too old? The impact of geochronology on cosmology, 1929-1952," in *The Age of the Earth: from 4004 BC to AD 2002*, ed. C.L.E. Lewis and S.J. Knell, 157-175. London: The Geological Society of London, 2001.

-Campbell, 1952. = Campbell, Norman. *What is Science?* New York: Dover Publications, 1952.

-Christianson, 1995. = Christianson, Gale E. *Edwin Hubble: Mariner of the Nebulae*. New York: Farrar, Straus, Giroux, 1995.

-Ellis, 1990. = Ellis, G.F.R. "Innovation, resistance and change: the transition to the expanding universe." In *Modern Cosmology in Retrospect*, ed. B. Bertotti, R. Balbinot, S. Bergia and A. Messina, 97-113. Cambridge: Cambride University Press, 1990.

-Fehrenbach, 1984. = Fehrenbach, Charles. "Twentieth-century Instrumentation." In *The General History of Astronomy*, Vol. 4, *Astrophysics and Twentieth-century Astronomy to 1950: Part A*, Owen Gingerich, ed., 166-189. Cambridge: Cambridge University Press, 1984.

-Galison, 1987. = Galison, Peter. *How Experiments End*. Chicago and London: University of Chicago Press, 1987.

-Garfield, 1985. = Garfield, Eugene. "The Life and Career of George Sarton: The Father of the History of Science." *Journal of the History of the Behavioral Sciences* 21 (April 1985): 107-117.

-Hetherington, 1982. = Hetherington, Norriss S. "Philosophical values and observation in Edwin Hubble's choice of a model of the universe." *Historical Studies in the Physical Sciences*, Vol. 13, Part 1 (1982), 41-67.

-Hetherington, 1988. = Hetherington, Norriss S. *Science and Objectivity: Episodes in the History of Astronomy*. Ames: Iowa State University Press, 1988.

-Hetherington, 1996. = Hetherington, Norriss S. *Hubble's Cosmology: A Guided Study of Selected Texts*. Tucson: Pachart Publishing House, 1996.

-Hoskin, 1976. = Hoskin, M. "The 'Great Debate': What Really Happened." *Journal of the History of Astronomy* 7(1976): 169 -182.

-Hoyt, 1980. = Hoyt, William Graves. "Vesto Melvin Slipher, November 11, 1875-November 8, 1969." *Biographical Memoirs of the National Academy of Sciences*, 1980, 52: 411-449.

-Jones, 1989. = Jones, Brian. "The Legacy of Edwin Hubble." Astronomy 17 (Dec. 1989): 38-44.

-Kragh, 1987. = Kragh, Helge. "The Beginning of the World: Georges Lemaître and the Expanding Universe." Centaurus, 1987, 32: 114-139.

-Mayall, 1970. = Mayall, N. U. "Edwin Powell Hubble: November 20, 1889 – September 28, 1953." *Biographical Memoirs of the National Academy of Sciences*, 1970, 41: 175-214.

-North, 1990a. = North, J.D. *The Measure of the Universe: A History of Modern Cosmology*. New York: Dover Publications, 1990.

-North, 1990b. = North, J.D. "The Early Years." In *Modern Cosmology in Retrospect*, ed. B. Bertotti, R. Balbinot, S. Bergia and A. Messina, 11-29. Cambridge: Cambride University Press, 1990.

-Osterbrock, 1990. = Osterbrock, Donald E. "The Observational Approach to Cosmology: U.S. observatories pre-World War II." In *Modern Cosmology in Retrospect*, ed. B. Bertotti, R. Balbinot, S. Bergia and A. Messina, 247-290. Cambridge: Cambride University Press, 1990.

-Sarton, 1927. = Sarton, George. *Introduction to the History of Science: From Homer to Omar Khayyam*. Baltimore: Williams and Wilkins, 1927.

-Sarton, 1931. = Sarton, George. *The History of Science and the New Humanism*. Bloomington: Indiana University Press, 1931.

-Sciama, 1971. = Sciama, D.W. *Modern Cosmology*. Cambridge: Cambride University Press, 1971.

-Seitter and Duerbeck, 1990. = Seitter, Walter C. and Hilmar W. Duerbeck. "Carl Wilhelm Wirtz: a pioneer in observational astronomy." In *Modern Cosmology in Retrospect*, ed. B. Bertotti, R. Balbinot, S. Bergia and A. Messina, 365-400. Cambridge: Cambride University Press, 1990.

-Shapin, 1994. = Shapin, Steven. *A Social History of Truth: Civility and Science in Seventeenth-century England*. Chicago: University of Chicago Press, 1994.

-Smith, 1982. = Smith, Robert W. *The Expanding Universe: Astronomy's "Great Debate" 1900-1931*. Cambridge: Cambridge University Press, 1982.

-Smith, 1990. = Smith, Robert W. "Edwin P. Hubble and The Transformation of Cosmology." Physics Today 43 (April 1990): 52-58.

-Zilsel, 2000. = Zilsel, Edgar. "The Genesis of the Concept of Scientific Progress and Scientific Cooperation." In *The Social Origins of Modern Science*. Dordrecht, Boston, and London: Kluwer Academic Publishers, 2000. A shortened version was published under the title "The Genesis of the Concept of Scientific Progress", *Journal of the History of Ideas*, 1945, 6, p. 325-349.